

Figure 1

CGAGAGACGACAGAAGGTTACGGCTGCGAGAAGACGACAGAAGGGTCCAGAAAAA
GGAAAGTGCTGGAGGGGAGTGGGGACAAAAGCAGCGACCAAGTGAATGTCACTTC
AGTGACTGAGGCCAGGCAAAACGCGCGGGAAGGATTTTGTGTAGCTTGGGACCCTT
TCATAGACACTGATGACACGTTTACGCAAAATAGAAATTTGAGGAGAAACGCCTGG
GCCTTCGGAAAGGAGTGATTGATTAGTACTTGCAAGTTTAGGTGACTTTAAGGAGAA
CTAACTAATGTATACTATTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGC
AGCAGGAAAGCTTTGGTTAATTTGGAAATGGATGATAGCATTAAAATAACAGAAGC
GCCTCCAGGTCTCTGAAGCTTCAGTCCCCCAGCTGAAAGCCAGAAAAGACTAAGCC
CACTAAGCCTTTTGATCCCTTTGGAAGCAAAGAACTTTCCTTCCCTGGGGTGAAGAC
TCTCCTCAGAAGATTTCTGTCTCTGCCTATGTTACAAGAGGAATCAAAACCAAGAC
AGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAAGTCAGCGGCTTGTTGATCTCCC
TCACTGCTGTCTGCCTGGTGGTCACCCCTGGGAGCAGGGCCTGTCCTCGCCGCTGTG
CCTGCTATGTGCCACAGAGGTGCACTGTACATTTTCGGTACCTGACCTCCATCCCAG
ATGGCATCCCGGCCAATGTGGAACGAATAAATTTAGGATATAACAGCCTTACTAGAT
TGACAGAAAACGACTTTGATGGCCTGAGCAAACCTGGAGTTACTCATGCTGCACAGT
AATGGCATTACAGAGTCAGTGACAAGACCTTCTCGGGCTTGACAGTCCTTGACGGTC
TAAAAATGAGCTATAACAAAGTCCAAATCATTTCGGAAGGATACTTTCTACGGACTC
GGGAGCTTGGTCCGGTTGCACCTGGATCACAACAACATTGAATTCATCAACCCTGAG
GCCTTTTATGGACTTACCTCGCTCCGCTTGGTACATTTAGAAGGAAACCGGCTCACA
AAGCTCCATCCAGACACATTTGTCTCATTAAAGCTATCTCCAGATATTTAAAACCTCTT
TCATTAAGTACCTGTTCTTGTCTGATAACTTCCTGACCTCCCTCCCAAAGAAATGGT
CTCCTACATGCCAAACCTAGAAAGCCTGTATTTGCATGGAAACCCATGGACCTGTGA
CTGCCATTTAAAGTGGTTGTCTGAGTGGATGCAGGGAAACCCAGATATAATAAAAT
GCAAGAAAGACAGAAGCTCTTCCAGTCCTCAGCAATGTCCCCTTTGCATGAACCCCA
GGATCTCTAAAGGCAGACCCTTTGCTATGGTACCATCTGGAGCTTTCCTATGTACAA
AGCCAACCATTGATCCATCACTGAAGTCAAAGAGCCTGGTTACTCAGGAGGACAAT
GGATCTGCCTCCACCTCACCTCAAGATTTTCATAGAACCCTTTGGCTCCTTGTCTTTGA
ACATGACANANNTNTCTGGAAATAAGGCCGACATGGTCTGTAGTATCCAAAAGCCA
TCAAGGACATCACCAACTGCATTCACTGAAGAAAATGACTACATCATGCTAAATGC
GTCATTTTCCACAAATCTTGTGTGCAGTGTAGATTATAATCACATCCAGCCAGTGTG
GCAACTTCTGGCTTTATACAGTGACTCTCCTCTGATACTAGAAAGGAAGCCCCAGCT
TACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGGTGGCTCTTAGGCCTGAAGA
CATTTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCCTTTTTGGTTCCAACAAGA
AAAAATTGTCTTGCAGCTGAACAGAACTGCCACCACACTTAGCACATTACAGATCCA
GTTTTCCACTGATGCTCAAATCGCTTTACCAAGGGCGGAGATGAGAGCGGAGAGAC
TCAAATGGACCATGATCCTGATGATGAACAATCCCAAACCTGGAACGCACTGTCCTGG
TTGGCGGCACTATTGCCCTGAGCTGTCCAGGCAAAGGCGACCCTTCACCTCACTTGG
AATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCCTTACGTTAGCGAGGATGGG

CGAATCCTAATAGACAAAAATGGGAAGTTGGAAGTGCAGATGGCTGACAGCTTTGA
TGCAGGTCTTTACCACTGCATAAGCACCAATGATGCAGATGCGGATGTTCTCACATA
CAGGATAACTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGC
ACACAGTGGTTACGGGTGAGACGCTCGACCTTCCATGCCTTTCCACGGGTGTTCCAG
ATGCTTCTATTAGCTGGATTCTTCCAGGGAACACTGTGTTCTCTCAGCCATCAAGAG
ACAGGCAAATTCTTAACAATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGAT
CAAGGTCATTACCAATGTGTGGCTGCCAACCCTCAGGGGCGGACTTTTCCAGTTTT
AAAGTTTCAGTTCAAAAGAAAGGCCAAAGGATGGTTGAGCATGACAGGGAGGCAG
GTGGATCTGGACTTGGAGAACCCAACCTCCAGTGTTTCCCTTAAGCAGCCAGCATCTT
TGAAACTCTCTGCATCAGCTTTGACAGGGTCAGAGGCTGGAAAACAAGTCTCCGGTG
TACATAGGAAGAACAACATAGAGACTTAATACATCGGCGGCGTGGGGAATTCCACG
CTCCGGCGATTACAGGGAGCATAGGAGGCAGCTCCCTCTCTCTGCTCGGAGAATTGAC
CCGCAACGCTGGGCAGCACTTCTAGAAAAAGCCAAAAAGAATTCTGTGCCAAAAAA
GCAAGAAAATACCACAGTAAAGCCAGTGCCACTGGCTGTTCCCCTCGTGGAAGTCA
CTGACGAGGAAAAGGATGCCTCTGGCATGATTCTCTCCAGATGAAGAATTCATGGTTC
TGAAAATAAGGCTTCTGGTGTCCAGGAAGGTACCAACTGCTGACTCTGGACCAG
TAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC
CACAAACACTACAATCTGAGCACCTTCCTGATTTCAAATTATTTAGTGTAACAAACG
GTACAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACA
ACCAACCAAAACCCAATCATTATCTTTCCATCAGTAGCTGAAATTCGAGATTCTGCT
CAGGCAGGAAGAGCATCTTCCCAAAGTGCACACCCTGTAACAGGGGGGAAACATGGC
TACCTATGGCCATACCAACACATATAGTAGCTTTACCAGCAAAGCCAGTACAGTCTT
GCAGCCAATAAATCCAACAGAAAGTTATGGACCTCAGATACCTATTACAGGAGTCA
GCAGACCTAGCAGTAGTGACATCTTCTCACACTACTGCAGACCCTAGCTTCTCCA
GTCACCCTTCAGGTTACACACCACTGCCTCGTCTTTATTTACATTCTAGAAACAA
CAATACAGGTAACCTTCCCCTTGTCAGGCACTTGGGAAGAGAGAGGACAATTTGGA
GCAGAGGGGAGAGTTAAAAACCCACATAGAACCCCAAGTTCTCCGACGGCATAGACAC
AGGACTGTGAGGCCAGCAATCAAGGGACCTGCTAACAAAAATGTGAGCCAAGTTCC
AGCCACAGAGTACCCTGGGATGTGCCACACATGTCCTTCCGCAGAGGGGGCTCACAG
TGGCTACTGCAGCACTGTCAGTTCCAAGTTCATCCACAGTGCCCTCCCCAAAACTA
ATAATGTTGGGGTCATAGCAGAAGAGTCTACCACTGTGGTCAAGAAACCACTGTTAC
TATTTAAGGACAAACAAAATGTAGATATTGAGATAATAACAACCACTACAAAATAT
TCCGGAGGGGAAAGTAACCACGTGATTCCTACGGAAGCAAGCATGACTTCTGCTCC
AACATCTGTATCCCTGGGGAAATCTCCTGTAGACAATAGTGGTCACCTGAGCATGCC
TGGGACCATCCAACTGGGAAAGATTGAGTGGAAACAACACCACTTCCCAGCCCCC
TCAGCACACCCTCAATACCAACAAGCACAAAATTCTCAAAGAGGAAAACTCCCTTG
CACCAGATCTTTGTAAATAACCAGAAGAAGGAGGGGATGTTAAAGAATCCATATCA
ATTCGGTTTACAAAAGAACCCAGCCGCAAAGCTTCCCAAATAGCTCCTCTTTTACC
CACAGGTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCCGCCACCAGC
TCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGTG

CCAGAAGTCTCTCAGCAGGGAAGAAGCAGCCCTTCACCAACTCCTCTCCAGTGCTTC
 CTAGCACCATAAGCAAGAGATCTAATACATTAAACTTCTTGTCAACGGAAACCCCCA
 CAGTGACAAGTCCTACTGCTACTGCATCTGTCAATTATGTCTGAAACCCAACGAACAA
 GATCCAAAGAAGCAAAAGACCAAATAAAGGGGCCTCGGAAGAACAGAAACAACGC
 AAACACCACCCCCAGGCAGGTTTCTGGCTATAGTGCATACTCAGCTCTAACAACAGC
 TGATACCCCCTTGGCTTTTCAGTCATTCCCCACGACAAGATGATGGTGGAAATGTAAG
 TGCAGTTGCTTATCACTCAACAACCTCTCTTCTGGCCATAACTGAACTGTTTGAGAA
 GTACACCCAGACTTTGGGAAATACAACAGCTTTGGAAACAACGTTGTTGAGCAAAT
 CACAGGAGAGTACCACAGTGAAAAGAGCCTCAGACACACCACCACCTCCTCAGC
 AGTGGGGGCGCCCCCAGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTT
 ACAGACAGCAAAGTCACATCAGCTTTCCAGATGACGTCAAATAGAGTGGTCACCAT
 ATATGAATCTTCAAGGCACAATACAGATCTGCAGCAACCCTCAGCAGAGGCTAGCC
 CCAATCCTGAGATCATAACTGGAACCACTGACTCTCCCTCTAATCTGTTTCCATCCAC
 TTCTGTGCCAGCACTAAGGGTAGATAAACCACAGAATTCTAAATGGAAGCCCTCTCC
 CTGGCCAGAACACAAATATCAGCTCAAGTCATACTCCGAAACCATTGAGAAGGGCA
 AAAGGCCAGCAGTAAGCATGTCCCCCACCTCAGCCTTCCAGAGGCCAGCACTCAT
 GCCTCACACTGGAATACACAGAAGCATGCAGAAAAGAGTGTTTTTGATAAGAAACC
 TGGTCAAAACCCAACCTTCCAAACATCTGCCTTACGTCTCTCTACCTAAGACTCTATTG
 AAAAAGCCAAGAATAATTGGAGGAAAGGCTGCAAGCTTTACAGTTCCAGCTAATTC
 AGACGTTTTTCTTCCCTTGTGAGGCTGTTGGAGACCCACTGCCCATCATCCACTGGACC
 AGAGTTTCATCAGGANTTGAAATATCCCAAGGGACACAGAAAAGCCGGTTCCACGT
 GCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTGGACAGTA
 CCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTTCATGTCTCTTTGTCTGTG
 GTTTTTTACCCGGCAAGGATTTTGGACAGACATGTCAAGGAGATCACAGTTCACTTT
 GGAAGTACTGTGGAATAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTC
 CTGGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAAGGGAAGCAGAAAGG
 TCTGGGTAACACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTG
 GTTTTTACAAGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCACTGTTGGTTAAGA
 TACAAGTCATCACAGCTCCCCCTGTCATTATAGAGCAAAAGAGGCAAGCCATCGTTG
 GGGTTTTAGGTGGAAGTTTGAAACTGCCCTGCACTGCAAAAGGAACTCCCCAGCCTA
 GTGTTCACTGGGTCCCTTATGATGGGACTGAACTAAAACCATTCAGTTGACTCATT
 CCAGATTTTTCTTGTATCCAAATGGAACCTCTGTATATAAGAAGCATCGCTCCTTCAGT
 GAGGGGCACTTATGAGTGCATTGCCACCAGCTCCTCAGGCTCAGAGAGAAGGGTAG
 TGATTCTTACTGTGGAAGAGGGAGAGACAATCCCCAGGATAGAAACTGCCTCTCAG
 AAATGGACTGAGGTGAATTTGGGTGAGAAATTACTACTGAACTGCTCAGCTACTGG
 GGATCCAAAGCCTAGAATAATCTGGAGGCTGCCATCCAAGGCTGTCATCGACCAGT
 GGCACAGAATGGGCAGCCGAATCCACGTCTACCCAAATGGATCCTTGGTGGTTGGG
 TCAGTGACGGAAAAAGACGCTGGTGACTACTTATGTGTGGCAAGAAACAAAATGGG
 AGATGACCTAGTCCTGATGCATGTCCGCCTGAGATTGACACCTGCCAAAATTGAACA
 GAAGCAGTATTTTAAGAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTTGACTGCA

AGGCCTCTGGCTCCCCTGTGCCTGAGGTATCCTGGAGTTTGCCTGATGGGACAGTGC
TCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCCTT
TTCCACAATGGAACCTTGTATTTCAACAACGTTGGGATGGCAGAGGAAGGAGATTAT
ATCTGCTCTGCCAGAACACCTTAGGGAAAGATGAGATGAAAGTCCACCTAACAGT
TCTAACAGCCATCCCACGGATAAGGCAAAGCTACAAGACCACCATGAGGCTCAGGG
CTGGAGAAACAGCTGTCCTTGACTGCGAGGTCACTGGGGAACCGAAGCCCAATGTA
TTTTGGTTGCTGCCTTCCAACAATGTCATTTCAATTCTCCAATGACAGGTTACATTTT
ATGCCAATAGAACTTTGTCCATCCATAAAGTGAAACCCTTGACTCTGGGGACTATG
TGTGCGTAGCTCAGAATCCTAGTGGGGATGACACTAAGACATACAACTGGACATT
GTCTCTAAACCTCCATTAATCAATGGCCTGTATGCAAACAAGACTGTTATTAAAGCC
ACAGCCATTTCGGCACTCCAAAAAATACTTTGACTGCAGAGCAGATGGGATCCCATCT
TCCCAGGTCACGTGGATTATGCCAGGCAATATTTTCTCCCAGCTCCATACTTTGGA
AGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAGATGAGGAACATCCGGCTTTC
TGACTCTGCGGACTTCACCTGTGTGGTTCGGAGCGAGGGAGGAGAGAGTGTGTTGGT
AGTGCAGTTAGAAGTCCTAGAAATGCTGAGAAGACCAACATTCAGAAACCCATTCA
ACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAACTGCTCTGTGGAT
GGGAACCCCCACCTGAAATTACCTGGATCTTACCTGACGGCACACAGTTTGCTAAC
AGACCACACAATTCCCCGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAA
GCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCTA
CATCGAGAAACTCATCCTGTTAGAGATTGGGCAGAAGCCAGTCATTCTGACATACGA
ACCAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTGTCTG
ATGGGATCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGAC
AGGCCTCAAGTGGATGGAAAATACATACTGCATGAAAATGGCACGCTGGTCATCAA
AGCAACAACAGCTCACGACCAAGGAAATTATATCTGTAGGGCTCAAAACAGTGTTG
GCCAGGCAGTTATTAGCGTGTGAGTGTGTTGTGGCCTACCCTCCCCGAATCATAA
ACTACCTACCCAGGAACATGCTCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGT
GTGGCCTTGGGAATCCCCAAGCCAAAAGTCACCTGGGAGACGCCAAGACACTCCCT
GCTCTCAAAGCAACAGCAAGAAAACCCCATAGAAGTGAGATGCTTCACCCACAAG
GTACGCTGGTCATTCAGAATCTCCAAACCTCGGATTCCGGAGTCTATAAGTGCAGAG
CTCAGAACCTACTTGGGACTGATTACGCAACAACCTTACATCCAGGTACTCTGACAGG
AAGGGGGAGACTAAAATTCAACAGAAGTCCACATCCACAGGGTTTATTTTTTGGAA
GAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGAATACATTTACAGT
ATTAAATTTACAATGGACATGCGATGAGACTTGTAATGAAAGCATTGTGAACTGA
AACCGAGTCTCTGTGGATCTCAAAGCAAACCTTAACTTAAGGCACTTTGATTTTGC
CAACAAATAATAACAAACATTAAGAGAAAAAAATGATCCACTACGAAATAACAAAC
GGCTAATGCACCTGAATTCTCAGTAAAAAGACCTTTCTCTCGCTAACAGTTGCCAGC
TGCCTCGTGTCTGTTTCTACCAATGTCACAAACATCGCACACAGGGTGAATGGAGT
CAACGGGAAAGATTAAGTTTGCAGTCTGTGTAATCTCAATGTACAAATATTCTGTC
NCTGGTTTATAAACATTTTGATAAAACCGAAAAA
AAAA (SEQ ID NO: 1)

Figure 2

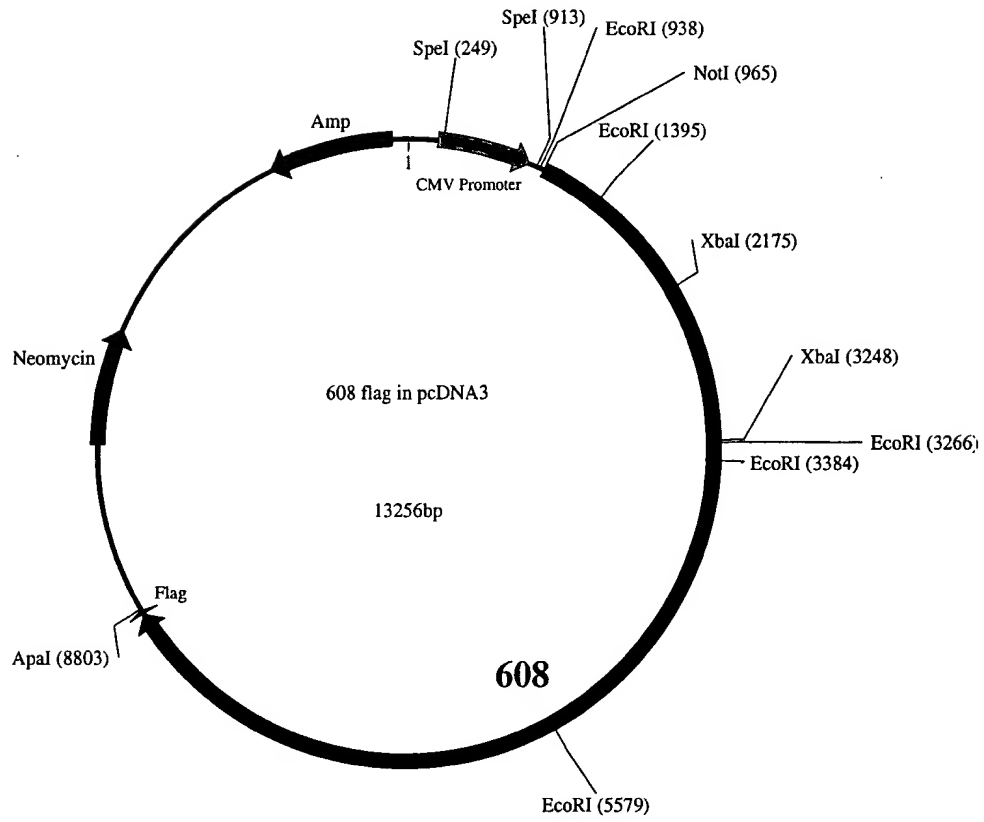


Figure 3

MQVRGREVSGLLISLTAVCLVVTTPGSRACPRRCACYVPTEVHCTFRYLTSIPDGIPANVE
 RINLGYNLTRLTENDFDGLSKLELLMLHSNGIHRVSDKTFSGLSLQVLKMSYNKVQII
 RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQI
 FKTSFIKYLFLSDNFLTSLPKEMVSYMPNLESYLHGNPWTCDCHLKWLSEWMQGNPDI
 IKCKKDRSSSSPQQCPLCMNPRISKGRPFAMVPSGAFLCTKPTIDPSLKSLSLVTQEDNGS
 ASTSPQDFIEPFGSLSLNMTXXSGNKADMVCSIQKPSRTSPTAFTEENDYIMLNASFSTNL
 VCSVDYNHIQPVWQLLALYSDSPLILERKPQLTETPSLSSRYKQVALRPEDIFTSIEADVR
 ADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWTMILMMNPNK
 LERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELQMA
 DSFDAGLYHCISTNDADADVLTyrITVVEPYGESTHDSGVQHTVVTGETLDLPCSTGV
 PDASISWILPGNTVFSQPSRDRQILNNGTLRLQVTPKDQGHYQCVAAANPSGADFSSFKV
 SVQKKGQRMVEHDREAGGSGLGEPNSSVSLKQPASLKLSSALTGSEAGKQVSGVHRK
 NKHRDLIHRRRGDSTLRRFREHRRQLPLSARRIDPQRWAALLEKAKKNSVPKKQENTTV
 KPVPLAVPLVELTDEEKDASGMIPPDEEFMVLKTKASGVPGRSPTADSGPVNHGFMTSI
 ASGTEVSTVNPQTLQSEHLPDFKLFSVTNGTAVTKSMNPSIASKIEDTTNQNPIIFPSVAE
 IRDSAQAGRASSQSAHPVTGGNMATYGHNTYSSFTSKASTVLQPINPTESYGPQIPITGV
 SRPSSSDISSHTTADPSFSSHPSGSHTTASSLFHIPRNNNTGNFPLSRHLGRERTIWSRGRV
 KNPHRTPVLRHRHRTVRPAIKGPANKNVSVQVPATEYPGMCHTCPSAEGTLVATAALS
 VPSSSHSALPKTNNVGVI AEESTTVVKKPLLLFKDKQNV DIEITTTTKYSGGESNHVIPTE
 ASMTSAPTSVSLGKSPVDNSGHL SMPGTIQTGKDSVETTPLPSPLSTPSIPTSTKFSKRKTP
 LHQIFVNNQKKEGMLKNPYQFGLQKNPAAKLPKIAPLLPTGQSSPSDSTLLTSPPPALST
 TMAATQNGKTEVVSGARSL SAGKKQPFTNSSPVLPSTISKRSNTLNFLSTETPTVTSPTAT
 ASVIMSETQRTSRKEAKDQIKGPRKNRNNANTTPRQVSGYSAYSALTADTPLAFSHSP
 RQDDGGNVSAVA YHSTTSLLAITELFEKYTQTLGNTTALETTLLSKSQESTTVKRASDTP
 PLLSSGAPPVPTSPPPFTKGVVTD SKVTS AFQMTSNRVVTIYESSRHNTDLQQPSAEAS
 PNPEITGTTDSPSNLFPSTSVPALRVDKPQNSKWKPSWPPEHKYQLKSYSETIEKGKRP
 VSMSPHLSLPEASTHASHWNTQKHA EKS VFDKKPGQNPTSKHLPYVSLPKTLLKKPRIIG
 GKAASFTVPANS DVFLPCEAVGDPLPIIHWTRVSSGXEISQGTQKSRFHVLPNGTLSIQRV
 SIQDRGQYLCSAFNPLGVDHFHVSLSVVFYPARILDRHVKEITVHFGSTVELKCRVEGMP
 RPTVSWILANQTVVSETAKGSRKVWVTPDGTLLIYNLSLYDRGFYKCVASNPSGQDSLL
 VKIQVITAPPVIEQKRQAIVGVLGGS LKLPCTAKGTPQPSVHWVLYDGT ELKPLQLTHS
 RFFLYPNGTLYIRSIAPSVRGTYECIATSSSGSERRVILTVEEGETIPRIETASQKWTEVN
 LGEKLLLNC SATGDPKPRIIWRLPSKAVIDQWHRMGSRIHVYPNGSLVVGSVTEKDAGD
 YLCVARNKMGGDDLVMHVRRLR LTPAKIEQKYFKKQVLHGKDFQVDCKASGSPVPEV
 SWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGTLYFN NVGMAEEGDYICSAQNTLGK
 DEMKVHLTVLTAIPRIRQSYKTTMRLRAGETA VLDCEVTGEPKPNVFWLLPSNNVISFS
 NDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSGDDTKTYKLDIVSKPPLINGLYANKT
 VIKATAIRHSKKYFDCRADGIPSSQVTWIMPGNIFLPAPYFGSRVTVHPNGTLEMRNIRLS

DSADFTCVVRSEGGESVLVVQLEVLEMLRRPTFRNPFNEKVIAQAGKPVALNCSVDGNP
PPEITWILPDGTQFANRPHNSPYLMAGNGSLILYKATRNLKSGKYRCAARNKVGYIEKLIL
LEIGQKPVILTYEPGMVKSVSGEPLSLHCVSDGIPKPNVKWTTTPGGHVIDRPQVDGKYIL
HENGTLVIKATTAHDQGNVICRAQNSVGQAVISVSVMVVA YPPRIINYLPRNMLRRTGE
AMQLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGTLVIQNLQTSDSGVY
KCRAQNLLGTDYATTYIQVL (SEQ ID NO: 2)

U.S. Patent & Trademark Office

Figure 4

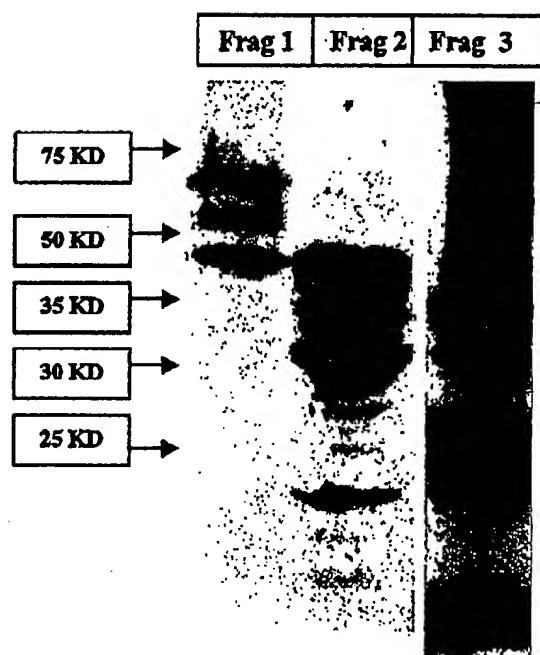


Figure 5

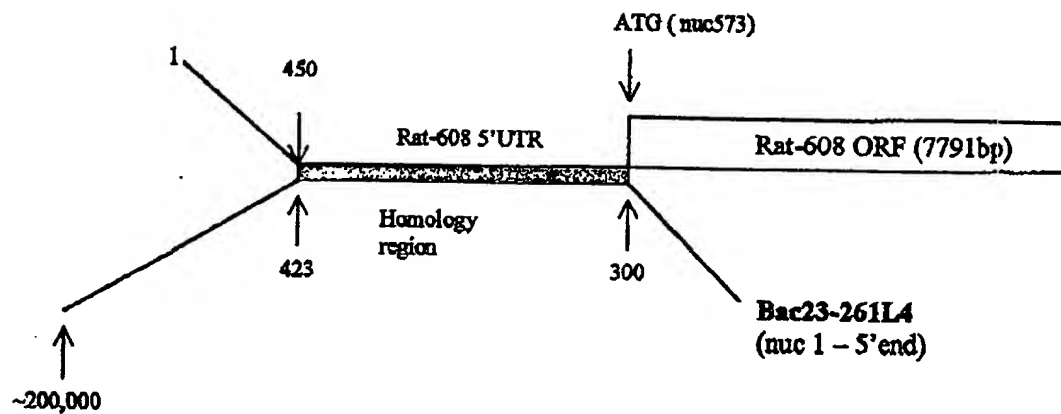
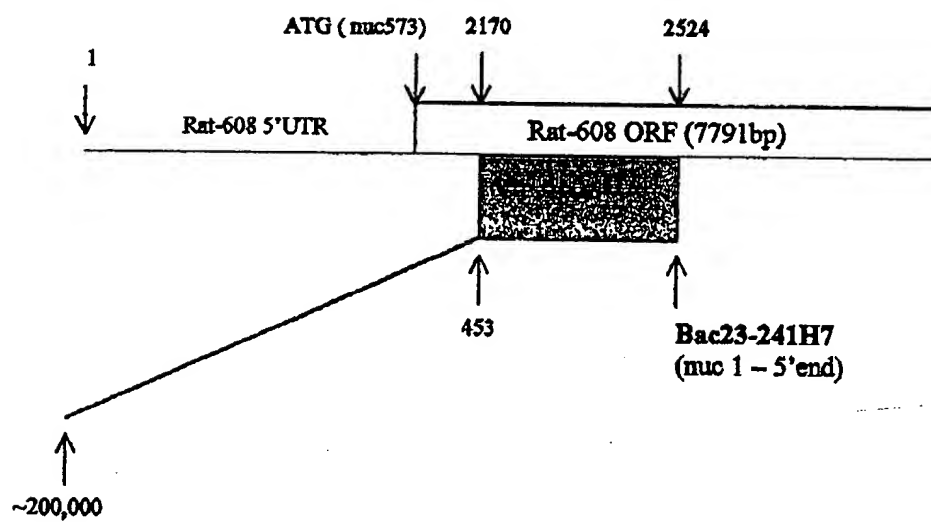


Figure 6



Nbc 1

↓ Nuc 390

▲ Nec 2817

+ Nov 2017
 AAAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTT
 GCAAACAGAAGGGGAGTCACAAACCCAGGACAAACCCACAGTGATTGAAGCGTCTTTGAA
 TGTATTGCTGTTGTTGTTACCAACATCAATTAGCATATATTCATTGTGAAAACCTTACGGGGT
 CTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGCACCACTACCA

▼ Exon 1 (Nuc 3661)

↑ Nuc 4203

TGTATAAATACAGTTTTGAGGCTTAAGTGTACGGGAACTCATGTGGTATTCAATTAACGGCTCT
 CTCTCTTATAACTAAGCTTAAGGTGCATATAGTCTCTCTGTTTCCAGCTACCTTGTAC
 CATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAATCATCACGCAGAGAGTATT
 CAAAAATATTCAGTGTATGTAACAGTGCAGGTGTAGGCATAGAAGTAATCATTAGTAAATC
 TTAATTTGGGTAAACTCATTCATAACAGCTOCAGGTGGGAGGGTCACTGAGGCTTGC
 CAGTGCGGGTAAAGATATTTTCTAACAAGAGAAGCAGAATTTCTCTCTGGCCATGCTCC
 CCATCACTGTGTCAGTAAGCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGT
 TGCTGCAAAAGTCAGAGAGCTCAGCCCTCCAGCTGGTCAGTTTACTGTCTCCCGGTCAAT
 AGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGCAGACATGCTAGAAA
 GAAATTTGACCTTTTCTAGATGGGTATTACAGCTGTAAAAGTATTTTGGAAAGGTAAAG
 CCAATAAATAAAGCATAATTAATAATCAATGTTACAAAAATTGATCATATAAAGAA
 GTACATTCATAAATGCAATGTGAAAAATATATATAAATTTTATCTATTACTGGTGCAAA
 TTTTCTAAATGTCAGTACTATTTTTATATTATAAAAAATTTTAAAAATGTATATAAA
 AGTGTAAAAAGGCTCTTGGTCAAAACAAGAGATTAAATTTACAAACTTTAATTGTCCCAT
 AACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGAAATGAGAAGCTATG
 AAGATATGTTTACAATAATAAGCCATTTAGTGATAAAGTCCAAATGGGAAGCTAGCACAC
 ACTGGTTTATAAAGAGAACAGTTTCTCTGAGTCTATGCAAGTTTACACTCTAGGGAAATAG
 AGTTCCTCTTCTCCAGATTTCACCTAGCATTTGTGTGCATCATTTTATCTCTGTGATGAG
 CATTATAAGTGGAAATAAGATAGGATCTCAAAGGAATGTCAATTTGGATGCCCTGAACAAT
 CTTTCAGGCTTTCTTTTCAGTTCACTAGTCTATTCATTTATTGGATAATTGGGGGATGGTGT
 TAATTTTTTTGCAGTCTTATGGAATTTCAAAAAACAACAAAAACAACAACAACA
 AACTCTGAAACTAGAACTACCAATCCATTACTGGGTATGTAAACAAAGAGAAATCTGCACA
 GAATTTATTGCTACATTTGTTTCATTATTCACACAGCCAGAATGTGGAACCAACTTACGTA
 GCGGTCAAAATATGAAAGGATAAAGAAAAATGTGGAATGTGTACAAACAGATGCCATGTG
 GCCATAAAGATGTGAAATCATGACATATGCAAGAAATGGATGTCAACTGGAAATCAATTT
 GGCTAATCAAAACAAGACAGAGCTCAAAAAGGAAACACCGGTGTAGCTTCTCTGACAAACA
 GAAGCTAGATTACACTTGTACGTGGCATGTGTGTTAGAAATTTATTAGTTATACACT
 ATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAGCAAAAAACAAGCTAGCTGGG
 TGGTAGGAGAGAAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAACATAGGTG
 CCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCAACAGACAGGAGGCCACAGTTCA
 AACAGGGTGAAATCCAGCACAGAGAAGGGGAAGTAGACACAAGTTTCCGCACTAAC
 CAAAGAAGCCATTTCAGTTGCTGCCTGCTGGGAGGGGCGTTCCAGTTTCTCCAGTCTGAC
 ACTGTGTATAACAACCAAGTTGACAATACAAAGTTGGCATGATGGATGGTTTTGTGCTATT
 TTTCATTTTCTTCTACGTGTTTGTGTGTGTGGTGGTGTGTGTGGTGGTGGCTGTGGTTTT
 ATTGTGTTCTTTGAGAGAGAGAAGGAACATGAAATGGGGTGGGTAGGAAGCTGGAAACG
 ATCTGGAAGAAAGTTGGGGAAAGAGAAAAATTGATGTGAGCATATTTAAACAAACAAACA
 AACAAACAAAGGTTCAATTTGCCACAAAGGTGTGAATTAATTAACAGTTACGACT
 CTTAAAGAAAATATTCCCAATTAATCCACAGATTGCTATGATGCTGTGGCTAGGACCTTGG

▼ Exon 2 (Nuc 6559)

CTTGAAC TGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGCAA
 GTCAGCTGCTGCTGATCTCCCTCACTGCCATCTGCTGGTGGTCACCCCTGGGAGCAGGGTC
 TGTCTCGCGATGTGCTGCTATGTGCCACAGAGGTGCACTGTACATTTGGGGACCTGADC
 TCCATCCAGACGGGCATCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGT

↑ Nuc 6755

GGACCTTGOCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCTGCTACTTTGCCCCCC
 AAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTTATAITG
 TGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAACTTGAAGCCCTCTTAGIT
 GTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTTAAATGAAAAAGAT
 ACAGTTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAATTGTCAATTGAATTTTAA
 AAAAGCATTTCTTTTGAATTTATAAATACCCATTACAAAATGGCTTACTTAAAATACTTG
 CTTTACTAAATCTGACAAATTATGGTGATATTTTGAAGGTTTATGAAAATTTGTTTATGT
 ATAAATGCACAAGAAATGGGATATGCCATCACCTATGTCCATTAGTGAGCATGTACAGT
 ATGCCAAACACTATTGTTTCAGTTTGGAGGAAGTAATGGGGGTGGGGGAGCAACAAGGGT
 TATAACCGTATACCCAGTGCCCTTGAAGCGATTGCAAAACAGTAAAGACTGACATTGTGTT
 CTCCCTATGAGGGAGGGGCTTGGGCTGAGCACTTTGCAATGAGCATTGTCTCATTGTGCT
 GGCAGGTTTTATGATAACTTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAACCTCAACT
 GAGAACATGCCTGAAGAAGATCAGATTATAGGCAGGCTGTGGGGCATTTTCTTAATTAG
 TGATTCATGGGGCAGGGCCAGTCCATTGTTCTGTGTACCATTCTCAGGCATTTAAAAA
 AAAAAAAAAACAGGCTGAGCAAGTGTCAAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGA
 TCTCTGCATCAGCTCCTGCCCTCCAGGTTCTACCTATTTGAGTTCCTGTCTAGCTCCCTA
 CAGTGATGAACAATGATGTGGAAAGTATAAGCCAAATAAATCCTTTCTTCCCAACTTGCTG
 TTGGTCATGATGTTTCATCAGTGATAATAGTCTCATGAAGATGCTGGTGTTTATAACA
 CCTTTGGACTAAATTTCTGTTATCTATAGCTGAGGAAAAATGGAGCATAGAAAGTCTCCAGA
 CTACACCAGAGTGTAACTGTGGGCTGAGCTTAGAATCACACCACGTGCCTCCACTGCC
 GGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCTGTTTGTTCATTTTGACA

Exon 3 (Nuc 8089) ▼

TGTGGACTTTAATTGACGATTTCATCTGAAGCTGAAAATGATTTTTTTCCAGGTATAACAGC
 CTCACTAGATTGACAGAAAATGACTTTTCTGGCTGAGCAGACTGGAGTTACTCATGCTGCACA
 GCAATGGCATTTCACAGAGTCAGTGACAAGACCTTCTGGGCTTGCAGTCTTGCAG

Nuc 8218 ↑

GTGAGATAGGTAGAGGGTGATGGAGGCTGAGAAGAGAGGTTGCAACTGTGGGTATACCC
 AAAAGCTGCTGATTCCCGTGGGAGACATTCTATAAGCATTCTATAAACTAGAGGCAGATA
 TCAAGGAAGGATTTCAATTGTAATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATG
 CTGGGGAAAATGCTTACACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAA
 AAAGCACAGCATGGCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTCAGAAT
 CCCAGAGACTCATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCACTGACAAA
 ACTTGACTCAGAACTAATGTGGAAGCATCAGGAAGACAGCCAACATCTGCTCTACT
 CATGCATGAATAAGGATCCAGAGAGAAAGGGAAGAAAAGGAAGGAAGGAAGGAAGG
 AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
 AAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
 GGAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGG
 TTTTATCCATAAAAGGTCAATTCACCTGTTTGTCTGGAAGTAGAGTGGGATCCCTTATAT
 AAGGGCAGTCTTTAACATAGTAGCATTTTATAAAACATTACAAATTTGAGTTTTCTCTAC
 TTTTATCTCTACCATCTTCAAACCTGAAACTACAATTATTOCCACAAATGAAGAAAATGC
 TGTAAAGATTTTACACACCCGAAGTGGGAACTTAAGGATTAGACAAGTCTAACAATGAG
 AATGGGGAGAACAAAAAGAGACTGCACAGGGAGCCCTTTCTCTGCTTATAATCTTGACAC
 TTGAGAAGCTAATTGACGCTGCATGACTACTCAACTCTTTAAGCAACAATGCTGTGTGTC
 ATGAAAAGCACAATAAAGTACATATGTCCATAATATTCATCAAAATTTGCATGCAGCAC
 ATAATAGCAATCAAAAGCAATAACACCCACTGTTTCACAGAGACTTTAAACATGAAACTGGA
 ACTATGTCTAGTGTGTTGACTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTG
 ATTTAGGTCATCAGACAGCATTGGAACATGTATCTTCAGGAGGAATCATTGATGTATCT
 GCATGAAATTTCTCCACCTATGTTTATTCTCTTAGCCAGGTTTTCTCTGATGGAGAAACATT
 GGGTTTGAGGTTTTACTCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGG
 CTTTATTATGAGGGATGTTGGTATTCCAGAAAATTTCTCTTTGAAGAGATTACAATTTA
 GGTCAAAAACAGAAAAATATGTAAAAAGTTATTGTTTTATTAAGTATTTCATGTTCTTTCTT
 TTTAAAAATGGTATGCTTAGAACTAATTAAGATTAGATTAGATTAGATTAGAAAATAATC
 AGAGAGGGATTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTTTGTCTCTAATT
 CAGAATCAATTAATTCATATTACTATAAAAGACAGCACGCCAGATGTGTGCCAGCTGAG
 GAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACAGAAAGGAGTGAAGGGTTGA

Exon 4 ? (Nuc 11286) ↓

Nuc 11680

Nuc 11967 ▲

(SEQ ID NO:3)

Figure 8

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length
1	1	542	542	2356
2	2898	3094	197	1334
3	4428	4557	130	3068
4	7625	8019	394	

Figure 9

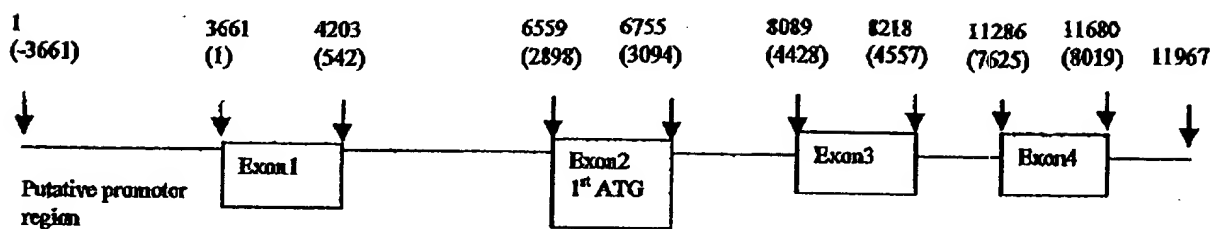


Figure 10

cDNA_rat	1	ogagagacgacagaagggttacggctgcgagaagaacagaagggtccag
genomic_hu	1	-----
cDNA_rat	51	aaaaaggaaagtgcctggaggggagtggggacaaaagcagcgaccaagtga
genomic_hu	1	-----
cDNA_rat	101	atgtcacttcagtgcactgaggccaggcaaaaacggcggaaggattttgt
genomic_hu	1	-----
cDNA_rat	151	gtagcttgggaocctttcatagacactgatgacacgtttacgaaaaatag
genomic_hu	1	-----
cDNA_rat	201	aaatttgaggagaaaacgcctgggccttoggaaggagtgattgattagta
genomic_hu	1	-----
cDNA_rat	251	cttgcagtttaggtgactttaaggagaactaactaatgtatactattga
genomic_hu	1	-----
cDNA_rat	301	gggaggaggaagagcattacagagtttccagcagcagcaggaaagctttg
genomic_hu	1	-----
cDNA_rat	351	gttaatttggaaatggatgatagcattaaaaataacagaagcgctccag
genomic_hu	1	-----
cDNA_rat	401	tctotgaagottoagtcctccagctgaaagccagaaaaagaotaagccac
genomic_hu	1	-----
cDNA_rat	451	taagccttttgatccctttggaagcaagaactttccttcctggggtga
genomic_hu	1	-----
cDNA_rat	501	agactctcctcagaagatttctgtctctgcctatgttacaagaggaatc
genomic_hu	1	-----
cDNA_rat	551	aaaaccaagacagaagagctcaggatgcaggtgagaggcagggaagtca
genomic_hu	1	-----
cDNA_rat	601	oggottgtgatctcctcctcactgctgtctgcctggtggtcaccctggga
genomic_hu	1	-----
cDNA_rat	651	gcagggcctgtcctgcgcgtgtgctgtatgtgccacagaggtgca
genomic_hu	1	-----
cDNA_rat	701	tgtacatttoggtaacctgacctccatccagatggcatccggccaatgt
genomic_hu	1	-----
cDNA_rat	751	ggaacgaataaatttaggatataacagccttactagattgacagaaaacg
genomic_hu	1	-----
cDNA_rat	801	actttgatggcctgagcaaaactggagttactcatgctgcacagtaatgg
genomic_hu	1	-----
cDNA_rat	851	attcacagagtcagtgacaagaccttctcgggcttgcaagtccctgcaggt
genomic_hu	1	-----
cDNA_rat	901	ottaaaaatgagotataaocaaagtccaaatoattoggaaggatactttct
genomic_hu	1	-----

cDNA_rat genomic_hu	951 1	acggactcgggagcttggtccggttgacactggatcacaacaacattgaa -----
cDNA_rat genomic_hu	1001 1	ttcatcaaccctgaggccttttatggacttacctcgctccgcttggtaca -----
cDNA_rat genomic_hu	1051 1	tttagaaggaaaccggctcacaagctccatccagacacatttgtotcat -----
cDNA_rat genomic_hu	1101 1	taagctatctccagatatttaaaccctctttcattaagtacctgttcttg -----
cDNA_rat genomic_hu	1151 1	tctgataacttcctgacctccctccaaaagaaatggtctcctacatgcc -----
cDNA_rat genomic_hu	1201 1	aaacotagaaagcctgtatttgcattggaacccatggacctgtgactgcc -----
cDNA_rat genomic_hu	1251 1	atttaaagtggttggtotgagtggtatgcagggaacccagatataataaa -----
cDNA_rat genomic_hu	1301 1	tgcaagaagagacagaagotcttccagtcctcagcaatgtccctttgcac -----
cDNA_rat genomic_hu	1351 1	gaaccccaggatctctaaaggcagaccctttgctatggtacccatctggag -----
cDNA_rat genomic_hu	1401 1	ctttcctatgtacaaagccaaccattgatccatcactgaagtcacagagcc -----
cDNA_rat genomic_hu	1451 1	ctggttactcaggaggacaatggatctgcctccacctcacctcaagattt -----
cDNA_rat genomic_hu	1501 1	catagaaaccccttggctccttctctttgaacatgacanannttctggaa -----
cDNA_rat genomic_hu	1551 1	ataaggccgacatggtctgtagtatccaaaagccatcaaggacatcacca -----
cDNA_rat genomic_hu	1601 1	actgcattcactgaagaaaatgactacatcatgctaaatgcgtcattttc -----
cDNA_rat genomic_hu	1651 1	cacaaatottgtgtgcagtgtagattataatcacatccagccagtggtgc -----
cDNA_rat genomic_hu	1701 1	aaottctggctttatacagtgactctcctctgatactagaaaggaagccc -----
cDNA_rat genomic_hu	1751 1	cagcttaccgagactccttcactgtcttctagatataaacaggtggctct -----
cDNA_rat genomic_hu	1801 1	taggcctgaagacattttaccagcatagaggctgatgtccagagcagacc -----
cDNA_rat genomic_hu	1851 1	ctttttggttccaacaagaaaaattgtcttgcagctgaacagaactgcc -----
cDNA_rat genomic_hu	1901 1	accacacttagoacattacagatccagttttccactgatgtctcaaatgcc -----
cDNA_rat genomic_hu	1951 1	tttaccaagggcgagatgagagcggagagactcaaatggacccatgatcc -----

cDNA_rat genomic_hu	2001 1	tgatgatgaacaatccccaaactggaaogcaactgttoctgggttgccggcact -----
cDNA_rat genomic_hu	2051 1	attgccotgagctgtccagggoaaaggogaccccttcacctcacttgggaatg -----
cDNA_rat genomic_hu	2101 1	gcttotagctgatgggagtaaagttagagcccccttacgttagcgaggatg -----
cDNA_rat genomic_hu	2151 1	ggcgaatcctaataagacaaaaatgggaagttggaactgcagatgggtgac -----
cDNA_rat genomic_hu	2201 1	agctttgatgcaggtotttaccactgcataagcaccaatgatgcagatgc -----
cDNA_rat genomic_hu	2251 1	ggatgtttcacatacaggataactgtggttagagccctatggagaaagca -----
cDNA_rat genomic_hu	2301 1	cacatgacagtgagggtccagcacacagtggttacgggtgagacgctcgao -----
cDNA_rat genomic_hu	2351 1	cttccatgcctttccacgggtgttccagatgottotattagctggattct -----
cDNA_rat genomic_hu	2401 1	tccagggaaacactgtgttctctcagcoatcaagagacaggcaaattctta -----
cDNA_rat genomic_hu	2451 1	acaatgggaccttaagaatattacaggttacgcacaaagatcaagggtcat -----
cDNA_rat genomic_hu	2501 1	taccaatgtgtgggtgccaaacooatoaggggcccgaactttocagtttta -----
cDNA_rat genomic_hu	2551 1	agtttcagttcaaaaagaaaggccaaaggatgggttagcatgcaggagg -----
cDNA_rat genomic_hu	2601 1	caggtggatctggacttggagaaocaaactccagtggtttcccttaagcag -----
cDNA_rat genomic_hu	2651 1	ccagcatctttgaaactctctgoatcagctttgacagggttoagagggtg -----
cDNA_rat genomic_hu	2701 1	aaaacaagtctccggtgtacataggaagaacaaacatagagacttantac -----
cDNA_rat genomic_hu	2751 1	atcggcggcggtggggattccacgctccggcgattcaggggagcataggagg -----
cDNA_rat genomic_hu	2801 1	cagctccctctctctgctcggagaattgaccgcgaacgctgggcagcact -----
cDNA_rat genomic_hu	2851 1	tctagaaaaagccaaaaagaattctgtgccaaaaaagcaagaaaatacca -----
cDNA_rat genomic_hu	2901 1	cagtaaaagccagtgccactgggtgttccctcgtggaactcactgaogag -----
cDNA_rat genomic_hu	2951 1	gaaaaggatgcctctggcatgattcctccagatgaagaattcatggttct -----

cdna_rat	3001	gaaaactaagggtcttctgggtgtccocaggaagggtcaccaactgtgactotg
genomic_hu	1	-----
cdna_rat	3051	gaccagtaaatcatgtggttttatgacgagtatagcttctggcacagaagtgc
genomic_hu	1	-----
cdna_rat	3101	tcaactgtgaatocacaaacactacaatctgagcaccttootgatttcna
genomic_hu	1	-----
cdna_rat	3151	attatttagtgtaacaaaacgggtacagctgtgacaaagagtatgaacccat
genomic_hu	1	-----
cdna_rat	3201	coatagcaagcaaaatagaagatacaacocaaaccccaatcattatc
genomic_hu	1	-----
cdna_rat	3251	tttccatcagtagctgaaattcgagattctgtctcaggcagggaagagcato
genomic_hu	1	-----
cdna_rat	3301	ttcccaaagtgcaacacocctgtaacagggggaaacatggctacctatggcc
genomic_hu	1	-----
cdna_rat	3351	atacoaacacatatagtagctttaccagcaaaagccagtagctcttgag
genomic_hu	1	-----
cdna_rat	3401	ccaataaatccaaacagaaagttatggacctcgatacctattacaggagt
genomic_hu	1	-----
cdna_rat	3451	cagcagacotagcagtagtgacatctcttctcacactactgcagacccta
genomic_hu	1	-----
cdna_rat	3501	gcttctcagtagtacccttcaggttcacacaccactgcctogtattttatt
genomic_hu	1	-----
cdna_rat	3551	caatttcctagaaacaacaatacaggttaacttccctgtccaggcactt
genomic_hu	1	-----
cdna_rat	3601	gggaagagagaggacaatttgagcagagggagaggttaaaaacccacata
genomic_hu	1	-----
cdna_rat	3651	gaaccccgagttotccgacggcatagacacaggactgtgaggccagcanc
genomic_hu	1	-----
cdna_rat	3701	aagggaacctgtacaacaaaatgtgagccaagttccagccacagagtacco
genomic_hu	1	-----
cdna_rat	3751	tgggatgtgccacacatgtccttccgagaggggtcaagtggtactg
genomic_hu	1	-----
cdna_rat	3801	cagcaactgtcagttccaagttcatccacagtgccctccccaaaactaat
genomic_hu	1	-----
cdna_rat	3851	aatgttggggtcacatagcagaagaggtctaccactgtggtcaagaaaccaot
genomic_hu	1	-----
cdna_rat	3901	gttactatttaaggacaaacaaaatgtagatattgagataataacaacca
genomic_hu	1	-----
cdna_rat	3951	ctacaaaatatccggaggggaaagtaaccaogtgattoctacggaagca
genomic_hu	1	-----
cdna_rat	4001	agcatgaacttctgtctcaacatctgtatccctggggaaatctcctgtaga
genomic_hu	1	-----
cdna_rat	4051	caatagtgggtcacctgagcatgcctgggaccatccaaactgggaaagatt
genomic_hu	1	-----
cdna_rat	4101	cagtggaacaaacaccacttcccagcccccctcagcacacccctcaataoaa
genomic_hu	1	-----

cDNA_rat genomic_hu	4151 1	acaagcacaaaattctcaaagaggaaaactccottgcaocagatctttgt -----
cDNA_rat genomic_hu	4201 1	aaataaccagaagaaggagggtgtttaagaatccatatcaattoggtt -----
cDNA_rat genomic_hu	4251 1	tacaaaagaaccagccgcaaaagottccaaaatagctcctottttacc -----
cDNA_rat genomic_hu	4301 1	acaggtcagagttcccccctcagattctacaactctcttgacaagtcgoc -----
cDNA_rat genomic_hu	4351 1	accagctctgtctacaacaatggotgccactcagaacaagggcactgaag -----
cDNA_rat genomic_hu	4401 1	tagtatacgggtgccagaagtctctcagcaggaagaagcagcccttoacc -----
cDNA_rat genomic_hu	4451 1	aaactcctotccagtgttctctagcaccataagoaagagatctaatacatt -----
cDNA_rat genomic_hu	4501 1	aaacttcttgtcaacggaaacccccacagtgacaagtcctactgctactg -----
cDNA_rat genomic_hu	4551 1	catctgtcattatgtctgaaacccaacgaacaagatccaaagaagcaaaa -----
cDNA_rat genomic_hu	4601 1	gacaaaataaaggggcctcggaagaacagaaaacaacgcaaacaccacccc -----
cDNA_rat genomic_hu	4651 1	caggoaggtttctggctatagtgcatactcagctctaacaacagotgata -----
cDNA_rat genomic_hu	4701 1	cccccttggttttcagtcattccccacgacaagatgatgggtggaatgta -----
cDNA_rat genomic_hu	4751 1	agtgoagttgcttatcaactcaacaacotctcttctggccataactgaact -----
cDNA_rat genomic_hu	4801 1	gtttgagaagtacaccagactttgggaataacaacagctttggaaacaa -----
cDNA_rat genomic_hu	4851 1	cgttggtgagcaaatcacaggagagtaaccacagtgaaaagagcctcagac -----
cDNA_rat genomic_hu	4901 1	acaccaccaccactcctcagcagtggggggccccagtgcccactccttc -----
cDNA_rat genomic_hu	4951 1	cccacctccttttactaagggtgtggttacagacagcaaaagtcacatcag -----
cDNA_rat genomic_hu	5001 1	ctttccagatgacgtcaaatagagtggtcaccatatatgaatcttcangg -----
cDNA_rat genomic_hu	5051 1	cacaatacagatctgcagcaaccctcagcagaggctagccccaatcctga -----
cDNA_rat genomic_hu	5101 1	gatcataactggaaccactgactctctccttaactctgtttccatccactt -----
cDNA_rat genomic_hu	5151 1	ctgtgccagcactaagggttagataaaccacagaattctaaatggaagccc -----
cDNA_rat genomic_hu	5201 1	tctccctggccagaacacaaatatcagctcaagtcatactccgaaccat -----

cDNA_rat	5251	tgagaagggc	5251	tgagaagggc
genomic_hu	1	-----	1	-----
cDNA_rat	5301	cagaggocagcactcatgcctcacactggaatacacagaagcatgcagaa	5301	cagaggocagcactcatgcctcacactggaatacacagaagcatgcagaa
genomic_hu	1	-----	1	-----
cDNA_rat	5351	aagagtgtttttgataagaaacctgggtcaaaaacccaacttccaaacatct	5351	aagagtgtttttgataagaaacctgggtcaaaaacccaacttccaaacatct
genomic_hu	1	-----	1	-----
cDNA_rat	5401	gccttacgtctctctacctaagactotattgaaaaagccaagaataattg	5401	gccttacgtctctctacctaagactotattgaaaaagccaagaataattg
genomic_hu	1	-----	1	-----
cDNA_rat	5451	gaggaaagggtgcaagctttacagttccagctaattcagaogtttttctt	5451	gaggaaagggtgcaagctttacagttccagctaattcagaogtttttctt
genomic_hu	1	-----	1	-----
cDNA_rat	5501	ccttgtagagggtgttgagagccactgccatcatccactggacacagagt	5501	ccttgtagagggtgttgagagccactgccatcatccactggacacagagt
genomic_hu	1	-----	1	-----
cDNA_rat	5551	ttcatcagganttgaaatatcccaagggacacagaaaagcoggttccacg	5551	ttcatcagganttgaaatatcccaagggacacagaaaagcoggttccacg
genomic_hu	1	-----	1	-----
cDNA_rat	5601	tgottocaaatggcaccttgtccatccagagggtcagtattcaggaccgt	5601	tgottocaaatggcaccttgtccatccagagggtcagtattcaggaccgt
genomic_hu	1	-----	1	-----
cDNA_rat	5651	ggacagtacctgtgtctgcatttaatacactgggcgtagaccattttca	5651	ggacagtacctgtgtctgcatttaatacactgggcgtagaccattttca
genomic_hu	1	-----	1	-----
cDNA_rat	5701	tgtototttgtctgtggtttttaccggcgaaggattttggacagacatg	5701	tgtototttgtctgtggtttttaccggcgaaggattttggacagacatg
genomic_hu	1	-----	1	-----
cDNA_rat	5751	tcaaggagaccacagttcactttggaagtactgtggaactsaagtgcaga	5751	tcaaggagaccacagttcactttggaagtactgtggaactsaagtgcaga
genomic_hu	1	-----	1	-----
cDNA_rat	5801	gtggagggtatgccaggcctacggtttctctggatacttgcaaaccaaac	5801	gtggagggtatgccaggcctacggtttctctggatacttgcaaaccaaac
genomic_hu	1	-----	1	-----
cDNA_rat	5851	ggtggtctcagaaacggccaagggaagcagaaaggtctgggtaacacctg	5851	ggtggtctcagaaacggccaagggaagcagaaaggtctgggtaacacctg
genomic_hu	1	-----	1	-----
cDNA_rat	5901	atggascattgatcatotataatctgagtcctttatgatcgtggtttttac	5901	atggascattgatcatotataatctgagtcctttatgatcgtggtttttac
genomic_hu	1	-----	1	-----
cDNA_rat	5951	aagtgtgtggccagcaaccatctggccaggattcactgttggttaagat	5951	aagtgtgtggccagcaaccatctggccaggattcactgttggttaagat
genomic_hu	1	-----	1	-----
cDNA_rat	6001	acaagtcacacagctcccctgtcattatagagcaaaagaggcaagcca	6001	acaagtcacacagctcccctgtcattatagagcaaaagaggcaagcca
genomic_hu	1	-----	1	-----
cDNA_rat	6051	tcgttggggttttagtggaagtttgaaaotgccctgcactgcaaaagga	6051	tcgttggggttttagtggaagtttgaaaotgccctgcactgcaaaagga
genomic_hu	1	-----	1	-----
cDNA_rat	6101	actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa	6101	actccccagcctagtgttcactgggtcctttatgatgggactgaactaaa
genomic_hu	1	-----	1	-----
cDNA_rat	6151	accattgcagttgactcattccagatttttctgtatocaaatggaactc	6151	accattgcagttgactcattccagatttttctgtatocaaatggaactc
genomic_hu	1	-----	1	-----
cDNA_rat	6201	tgtatataagaagcatcgctccttcagtgggggcacttatgagtgcatt	6201	tgtatataagaagcatcgctccttcagtgggggcacttatgagtgcatt
genomic_hu	1	-----	1	-----
cDNA_rat	6251	gccaccagctcctcagggtcagagagaagggtagtgttcttactgtgga	6251	gccaccagctcctcagggtcagagagaagggtagtgttcttactgtgga
genomic_hu	1	-----	1	-----

ctg	rat	7351	CttTGGAAGCAGAGTCACGGTCCATCCAAATGGAACCTTGGAgatgagga
---	le_hu	890	CtaTGGAAGCAGAATCACAGTCCATAAAATGGAACCTTGGAsattagga
gat	rat	7401	acatccGGCTTTCTGACTCTGCCGACTTCACCTGTGTGGttcggagcgag
---	le_hu	940	atgtgaGGCTTTTACAGATTCAGCCGACTTTATCTGTGTGGcccgaaatgaa
oca	rat	7451	ggaGGAGAGAGTGTGTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAG
---	le_hu	990	ggtGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAATGCTGAG
IGg	rat	7501	AAGACCAACATTTCAGAAACCCATTCAACGAAAagtcacgcgccaagctg
IGt	le_hu	1040	AAGACCGACATTTAGAAATCCATTTAATGAAAAaatagttgccagctgg
3CA	rat	7551	gcaagcccgtagCACTGAAGTGTCTGTGGATGGGAACCCCCCACCCTGAA
3CA	le_hu	1090	gaaagtccacAGCATTGAATTGCTCTGTGATGGTAACCCACCACCTGAA
ATT	rat	7601	attacCTGGATCTTACCTGACGGCACACAGTTTGCTAACAGACCAacaaa
ACT	le_hu	1140	ataatCTGGATTTTACCAAATGGCACACGATTTTCCAATGGACCAcaaaag
CC	rat	7651	ttccccGTATCTGATGGCAGGCAATGGCTCTCTCATCCTTTACAAAgcaa
CC	le_hu	1190	ttatcagTATCTGATAGCAAGCAATGGTTCTTTTATCATTTCTAAAacaa
CT	rat	7701	ctcggaacaagtCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTGGC
CA	le_hu	1240	ctcgggaggatgCAGGAAAATATCGCTGTGCAGCTAGGAATAAGGTGGC
CA	rat	7751	TAcacgagaaactcatcctgttagagATTGGGCAGAAGCCAGTCATTCT
CA	le_hu	1290	TAtattgagaaattagtoaatattagaaATTGGCCAGAAGCCAGTTATTCT
TG	rat	7801	GACATACGAACACAGGGatggtgaagagcgTCAGTGGGAACCGTTATCAC
TG	le_hu	1340	TACCTATGCACCAGGGAacagtaaaaggoatCAGTGGAGAATCTCTATCAC
TC	rat	7851	TGCATTGTGTGTCTGATGGgATCCCCAAGCCAAATGTCAAGTGGACTACA
CT	le_hu	1390	TGCATTGTGTGTCTGATGGaATCCCTAAGCCAAATATCAAATGGACTATG
AC	rat	7901	CCGGGTGGccatGTAAATCGACAGGCCTCAAGTGGATGGAAAATACATAcT
AC	le_hu	1440	CCAAGTGGttatGTAGTAGACAGGCCTCAAATTAATGGGAAATACATatT
GA	rat	7951	GCATGAAAATGGCAGCTGGTCATCAAAGCAACAACAGCtcacgaccaAG
GA	le_hu	1490	GCATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCttatgacagAG
aa	rat	8001	GAAATTATATCTGTAGGGCTCAAAACAGTGTGGCCAggcagttattagc
at	le_hu	1540	GAAACTATATCTGTAAGGCTCAAAATAGTGTGGTCatacaotgattact
TT	rat	8051	gtgtCAGTGATGGTTGTGGCCTACCCCTCCCGAATCATAAActacotACC
TT	le_hu	1590	gttcCAGTAATGATTGTAGCCTACCCCTCCCGAATTACAAAtogtccACC
TA	rat	8101	CAGGAACATGCTCAGGAGGACAGGGGAAGCCatgCAGCTCCACTGTGTGG
TA	le_hu	1640	CAGGAGTATTGTCACCAGGACAGGGGCAGCctttCAGCTCCACTGTGTGG
TC	rat	8151	CCTTGGGAATCCCCAAGCCaaAAGTCACCTGGGAGACGCCAAGACACTCC
C	le_hu	1690	CCTTGGGAGTTCCCAAGCCagAAATCACATGGGAGATGCCTGACCACTCC
C	rat	8201	CTGCTCTCAAaagcaacagcaagaaaaccCATAGAAGTGAGATGCTTCA
C	le_hu	1740	CTTCTCTCAAcggcaagttaaagagaggacaCATGGAAGTGAGCAGCTTCA
TT	rat	8251	CCCACAAGGTACGCTgGTCATTTCAGAATCTCCAAACCTCGGATTCCGGag
TT	le_hu	1790	CTTACAAGGTACCCTaGTCATTTCAGAATCCCCAAACCTCCGATTCTGGga
TT	rat	8301	tcTATAAGTGACAGAGCTCAGAACCTACTTGGgactGATTACGCAACAACT
TT	le_hu	1840	taTACAAATGCACAGCAAAGAACCCACTTGGtagTGATTATGCAGCAACG
TT	rat	8351	TACATCCAGGTACTCTGACAGGAagggggagactaaaattoaacagaaagt
TT	le_hu	1890	TATATTCAGTAATCTGACATGAaataataaagtcaacaacatctgggca

cDNA_rat	8401	ccacatccacagggTTTATTTTTTGAAGAAGTTTAATCAAAGGCAGCCA
genomic_hu	1940	gaa-----TTTATTTTTTGAAGAAGTTTAATCAAAGGCAGCCA
cDNA_rat	8451	TAGGCATGTAAATGAgctcGAATACATTACAGTATTAAATTTACAATGG
genomic_hu	1979	TAGGCATGTAAATGAatttGAATACATTACAGTATTAAATTTACAATGA
cDNA_rat	8501	ACATGCgatga-----GACTTGTAATGAAAGCATTGTGAAGTGAaaccg
genomic_hu	2029	ACATGCaaaataaaaagGACTTGTAATAAATGCATTATGAAGTGAatgata
cDNA_rat	8546	agtctctg-----TGGATCTCAAAGCAAACCTCTTAAGTGAAGGCACTTTg
genomic_hu	2079	ctgatttatttaaTGGATCTCAAACAAACTTTTAACTTAAGGCACTTTt
cDNA_rat	8591	ATTTTGCCAACAAATAATAACAAACattaagagaaaaaaatgatcCACTA
genomic_hu	2129	ATTTTGCCAACAAATAACAATAAACaaacattgaaacggtt----CACTA
cDNA_rat	8641	CGAAATAACAAACGGCTAATGCACCTGAATTtctcagtaaaaagacotttc
genomic_hu	2175	TAAATAACAAATGGCTAATGTACCTGAATTtttcagtasaaaaaatgaac
cDNA_rat	8691	tctcgctaacagttgCCAGCTGCCTCGTGTCTGTTTCCCTACCAATGTCAC
genomic_hu	2225	ttctaata-----CCAGTTGCCTAGTGTCCACCTCCTATCAATGTTAC
cDNA_rat	8741	AAACatogcacacagggTGAATGGAGTCAACGGGAAAGATTAAGTTTGCG
genomic_hu	2268	AAGCATggcactca----GAACAGAGACAATGGAAATATTAATCTGCR
cDNA_rat	8791	GTCTgtgtaaatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgatgtaaatattaccatcctgATGTATAAATATTTTG----TG

exon1 (2342-2397)

cDNA_rat	8791	GTCTgtgtaaatctca-----ATGTACAAATATTCTGtcncTG
genomic_hu	2314	ATCtttatgatgtaaatattaccatcctgATGTATAAATATTTTG----TG
cDNA_rat	8829	GTTTATAAACATTTTGATAAAACCGAAAAAAAAAAAAAAAAAAAAA
genomic_hu	2360	GTTTATAAATTTTTTGCTAAACCTACAGAAATAAGcactgaactgtc

(Genomic human OCP: SEQ ID NO:4)
(cDNA rat: SEQ ID NO:5)

Figure 11

Exon/Intron No.	Exon start	Exon end	Exon length	Intron length	Remarks
1	1	208	208	69	No valid splice site found upstream this exon
2	277	429	153	18	
3	447	485	39	1561	
4	2046	2244	199	1351	
5	3595	3724	130	3254	
6	6978	7359	382	4123	
7	11482	14903	3422	38	
8	14941	15307	367	51	
9	15358	15825	468	1039	
10	16864	~17760	~897	?	Last exon might be complete
11	?(1)	2317	~231 7	25	Exon is not complete and start site is not known
12	2342	2397	56	-	

[illegible]

ATGAGAGGTTAAAGAGGCAAGGAACTACCTGCTTGTGCTGGTCTCCCTTTGCTGT
GATCTGCCTGGTTCGCCACCCCTGGGGGCAAGGCTGTCTCGCCGCTGTG
CCTGTTATATGCTACGGAGGTACACTGCACATTTCCGTTACCTGACTTCCA
TCCCAGACAGCATCCCCGCCAATGTGGAACGCATCAATTTAGGATAACAAC
AGCTTGGTATGATTGATGGAACAGATTTTCTGGCCTGACCAAACTGGA
GTACTCATGCTTCACAGCAATGGCATTACACAAATCCCTGACAAGACCTT
CTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTATAATAAAGTCC
GAAAACTTCAGAAAGATACITTTTATGGCCTCAGGAGCTTGACACGATTG
CACATGGACCAACAATAATTTGAGTTTATAAACCCAGAGGTTTTTATGG
GCTCAACTTTCTCCGCTGGTGCACTTGAAGGAAATCAGCTCACTAAGCT
CCACCCAGATACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTC
TTTCATTAAGTTCCTATACTGTCTGATAACTCTCTGACCTCCCTCCCTCAA
GAGATGGTCTCCTATATGCCTGACCTAGACAGCCTTTACCTGCATGGAAA
CCCATTGGACCTTGATTGCCATTTAAAGTGGTTGTCTGACTGGATACAGCC
AGATGTAAATAAAATGCAAAAAAGATAGAAGTCCCTCTAGTGTCTCAGCAGT
GTCCACTTTGCATGAACCTTAGGACTTCTAAAGGCAAGCCGTAGCTATG
GTCTCAGCTGCAGCTTTCAGTGTGCCAAGCCAACCATTGACTCATCCCTG
AAATCAAAGGCTGACTATCTTGGGAAGCAGTAGTTCTGTCTTTCATCTCT
CCCCAAGGTTTCTAGCACCCTTGTGGTCCCTCACTTTGAATGTACAGT
CAGTCTGGAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAG
GACATCACCCATTGCATTCAGTGAAGAAAATGACTACATCGTGCTAAATA
CTTCATTTTCAACATTTTTTGGTGTGCAACATAGATTACGGTCACATTACG
CAGTGTGGCAAATTTTGGCTTTGTACAGTATTCTCCTCTGATACTAGAAA
GGAGCCACTTGCTTAGTGGAACACCCGAGCTCTATTACAAATATAAACAG
GTGGCTCCTAAGCCTGAAGACATTTTTACCAACATAGAGGCAGATCTCAG
AGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAGCTGAACA
GAACTGCCACCAATTCAGTAGCATTACAGATCCGACTCTCCAGTGATGCT
CAAACTCACTTTACCAAGACAGAGTACAGGCGAGTGAACACAAACATGGT
CTATGATTTCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTA
GGTGGAACCGTTGGCCTGAACTGCCAGGCCAAGGAGACCCCAACCCACA
CGTGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTCA
GTGAGGATGGACGGATCCTAATAGACAAAAAGTGGAATTTGGAACCTCA
GATGGCTGATAGTTTGCACAGGCGTATATCACTGTATAAGCAGCAATTT
ATGATGATGCAGATATTTCTACCTATAGGATAACTGTGGTGAAGACCTTTG
TCGAAGCCTATCAGGAAAAATGGGATTTCACACACAGTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTTCTACTGGTATCCAGATGCTCTATTAGC
TGGGTTATTCCAGGAAACAATGTGCTCTACGTCAATCAAGAGACAAGAA
AGTTCTAAACATGGCACATTAAGAATAATTACGGTCACCCCGAAAGACC
AAGGTTAATCTCGCTGTGTGGCAGCCAACCCATCAGGGGTTGATTTTTGA
TTTTCCAAGTTTCAGTCAAGATGAAAGGACAAAGGCCCTTGGAGCATGAT
GGAGAAACAGAGGGATCTGGACTTGATGAGTCCAATCCTATTGCTCATCT
TAAGGAGCCACCAAGGTGCACAACCTCCGTACATCTGCTGTGATGGAGGCTG
AAGTTGGA AAAACACACCTCAAGCAACAGTAAGAGGCACAACTATCGGGA
ATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGA
ATAGGAGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGG
GCGGCATCTGTGGAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAG
AAAATACCACAGTAGGCCACCCCAAGGGTGGTCAACCAACTCCCAACATA
CCTGGTGAAGAAGACGATTCTCTCAGGCATGCTCGCTCTACATGAGGAATT

TATGGTCCCGGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGA CTGACTCCAGAACAATATCTGATAGTCCATGACAAACATAAAATTATGGC ACAGAAATCTCTCCTGTTGTGAATTCACAAATACTACCACCTGAAGAACCC ACAGATTTCAAACGTGTCTACTGCTATTAATAACTACAGCCATGTCAAAGAA TATAAACCCCAACCATGTCAAGCCCAAAATACAAGGCACCAACCAACATT CATCCACTGTCTTTCCACTGTCTACTTTGGAGCAACTGAAATTTCCAGGACTCTG ACCAGATGGGAAGAGGAAGAGAGCATTTCCAAAGTAGACCCCCAATAAC AGTAAGGACTATGATCAAAGATGTCAATGTCAAAATGCTTAGTAGACACCA CCAACAAACATTAATTAGAGTTCAGTAATAATACCACAAATAGTCATGACACA TCTGTAAGAGAAGATGTAGTGAACCCAGGCACAAATCACTTCTATTCTCACAC TACTCAAATACCTTAGCACCTCCACGTTCCCTTCAGATCCACACACAGCTGC TCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTT CAGACGCTTTGGGAGGCAGAGGAAAAATTTGGCGGAAGGGGGCGGATTATC AGCCCCATATAGAACAATCCAGTTCTGCGACGGCATAGATACAGCATTTTTCG GTCAACAACACGAGGTTCTTCTGTAAAAAAGCACTACTGCAATTCACGCCA CAGTGCTCAATGTGACATGTCTGTCTCTGCTTTCCAGGGAGAGGCTCACCA CTGCCACAGCAGCATTGTCTTTTCCAAGTGCTGCTCCCATCACCTTCCCCA AAGCTTGACATTGTAGAGTCCCATCAGAAGAGTCTACAACCTCTAGTCCAG AATCCCACTATTACTACTTGTAGAACAACCCAGGTAGAGAAAAACAACCC CACAATAAAAAATTTCCAGGACTGAAATTTCCCAAGTGAGTCCAACCTGGTG CAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAACTCACAAA GTAAACGCCAGTTACCCACGTGTGTCTAGACCAATGAAGCTAAAAGAGA TTCAGTGATTACATCGTCACTTTCAGGTGCTATACCAAGCCACCAATGAC TATTATAGCCATTACAAGTTTTCAGAAGGAAATTTCCCTGGCAACAGA ACTTTGTAATAAACCATACAACCAAAAGGCAGATTAAAGGAATCAACATAAA GTTAGTTTACAAAAAAGCACAGCTGTGATGCTTCTTAAACATCTCCTGCT TTACCACAGAGACAAAGTTCCCTTTCCATTTACCACACTTTCAACAAGT GTGATGCAAATTCATCTAATACCTTGACTACCGCTCACCACACTACGACC AAAACACACAATCTTGGAAGTCTTCCAACAAGAAGGAGCTTCCCTTCCC ACCCTTAACCCATAGTGCTTCTAGTATTATAAGCAAGAGCTCAAGTACAA AAAGCATCATATCAACGCAAAACAGCAATACCAGCAACAACCTCTACCTTC CCTGCATCTGTATCACTTATGAAACCCAAACAGAGAGATCTAGAGGACA AACAAATCAAAAGAGAACAGGAGCCTCAAAAAGAAACAGGACTGACCCA AACATCTCTCCAGACGACAGGATTCTGGCTTCTACTACACCCATGCTATGACA CCTCTGCTCTGGCATTCACTATTTCCCAACAGAAAAACAACACTGGGATT TCAAGCACAAATCAGTTTTTCATTCAAGAACTCTTAATCTGACAGATGTGATT GAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAATTGCTTC TGAACAACAACTTTGTCCAGCAAAATCACCCAGAGTACCACAACCTAGGAAG CATCATTTAGACACTCCCATACCACCTTCTTGAGCAGCAGTGTACTCTAA TGGCAGTTCCCATCTCCCCCTCCCTTTACTCAGAGAGCAGTTACTGACACAG GTGGCGACTCCCATTTCGGCTTATGACAAATACAGTGGTCAAGCTGCAC GAATCCTCAAGGCACAATCTCCAATGCCAAGTTCACAAATTGGAACCACT CACTTCACTTACCTCTAATCTGTTACATTTACTCTCCATGCCAGCATAAC AACAGTTAAATCACAGAAATCCAAATTAACCTCATCTCCCTGGGCAGAAT ACCAATTTTGGCACAAACCATACTCAGACATTGCTGAAAAGGCCAAAAAG CCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCT TGTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTTTGATA AGAAACCAAGTTCAAGAAGCAACAACCTCCAACCTTCCCTTTGACTCTT TGCTTAGGTATATATTTGAAAAGCCCAAGTATCTTGAGGAGAAAAGCTGCA AGTTTACTAATTCCAGCTAACTCAGATGGCTTTTCTCCCTGTGAAGCTGTT

GGAAATCCCCTGCCACCACCATTCATTGGACAGCAGAGTTTCAGGACTTGATTATTA
TCTAGAGGAAACCAGAATAGCAGGGTCCAGGTTCTCCCCAATGGTACCCT
GTCCATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTACTTGTGTTCCG
CATCCAATCTGTTTGGCACAGACCACCTTCATGTCACCTTGTCTGTGGTTT
CCTATCTCTCCAGGATCCTGGAGAGACGTCACAAAGAGATCACAGTTTCAT
TCCGGAAGCAGTGTGGAACCTGAAGTGCAGAGCAGAAGGTAGGCCAAGCC
CTACAGTTACCTGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCC
AGGGAAGTAGGCAGGCTGTGGTGACGGTTGACGGAACATTGGTCCTCCAC
AATCTCAGTATTATGACCGTGGCTTTTACAAATGTGTGGCCAGCAACCCA
GGTGGCCAGGATTCAGTCTGGTTAAAAATCAAGTCATTGACGACCACC
TGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGCACCTTGGGGTGAAA
GTTTAAACTGCCCTGTACTGCAAAGGAACTCCTCAGCCCAGCGTTTACT
GGGTCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCCA
AGTTGTTCTTATTTTCAAATGGGACCTTTGTATATAAGAAACCTAGCCTCTT
CAGACAGGGGCAATTGAAATGCATTGTCTACAGTTCCACTGGTTCCGGAG
CGAAGAGTGAGTAATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGA
TAGAAGCTGCATCCCAGAAAAAGGACTGAAGTGAATTTGGGGACAAATTA
CTACTGAACTGCTCAGCCACTGGGGAGCCCAAACCCCAATAATGTGGAG
GTTACCATCCAAGGCTGTGGTCGACGAGTGAGCTGGATCCAGCTTACC
CTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGACAGTGGTGTC
TACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATGCA
TGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTA
GAAAGCAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCC
GGCTCCCCAGTGCCAGAGATATCTTGGAGTTTGCTGATGGAACCATGAT
CAACAATGCAATGCAAGCCGATGACAGTGCCACAGGACTAGGAGATAT
ACCTTTTCAACAATGGAACCTTTATACTTCAACAAAGTTGGGGTAGCGGA
GGAAGGAGATTATACTTGCTATGCCCAGAACACCTAGGGAAAGATGAA
ATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGAATAAGGCAGAG
TAACAAAACCAACAGAGAATCAAAGCTGGAGACACAGCTGTCTTGACT
GTGAGGTCACTGGGGATCCCAACAAAAAATTTTGGTTGCTGCCTTCC
AATGACATGATTTCTTCTCCATTGATAGGTACACATTCATGCCAATGGG
TCTTTGACCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTACGTATG
TGTAGCCCCGAAATCCCAGTGGGGATGACACCAAAATGTACAAACTGGATG
TGGTCTTAAACCTCCATTATCAATCGTCTGTATACAAACAGAACTGTTA
TTAAAGCCACAGCTGTGAGACATTCCAAAAAACACTTTGACTGCAGAGCT
GAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCAGACAATATTTT
CCTCAGACCCCATACTATGGAAGCAGAAATCACAGTCCATAAAAAATGGAA
CCTTGGAATATTAGGAATGTGAGGCTTTTCAGATTACGGCCAGCTTATCTGTG
TGGCCCGAAATGAAGGTGGAGAGCGTGTTGGTAGTACAGTTAGAAAGTA
CTGGAATGCTGAGAAAGACCGACATTTAGAAATCCATTTAATGAAAAAT
AGTTGCCAGCTGGGAAAGTCCACAGCATTGAATTGCTCTGTTGATGGTA
ACCCACCACCTGAAATAATCTGGATTTTACCAAATGGCACAGATTTTCCA
ATGGACCACAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTATCA
TTTCTAAAAACACTCGGGAGGATGCAAGGAAATATCGCTGTGCGAGCTAGG
AATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATTGGCCAGAA
GCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGAG
AATCTCTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCA
AATGGACTATGCCAAGTGGTTATGTAGTAGACAGCGCTCAAATTAATGGG
AATAACATATTGCATGACAATGGCACCTTAGTCATTAAGAAAGCAACAGC
TTATGACAGAGGAAACTATATCTGTAAGGCTCAAAATAGTGTGGTCATA

Figure 13

A

Region			Region Length	% identity	% positives	% gaps
General	Rat	Human				
1-655	1-655	1-653	655	76	86	0
656-726	656-726	654-724	71	46	62	0
727-779	727-779	725-777	53	77	86	0
780-1634	780-1617	778-1612	655	38	53	4
1635-end	1618-end	1613-end	980	74	85	0
	Total	Alignment	2614	62	74	1

B

Region			% identity	% positives	% gaps
Rat	Mouse	Length			
1-238	1-238	238	91	92	1

C

Region			Region Length	% identity
General	Rat	Human		
1-1965	1-1965	1-1965	1965	83
1966-2178	1966-2178	1966-2178	213	86
2179-2337	2179-2337	2179-2337	159	86
2338-4893	2338-4893	2338-4863	2565	63
4894-7833	4852-7791	4864-7761	2940	80
	Total	Alignment	7842	80

D

Region			Region Length	% identity
General	Rat	Mouse		
1-720	1-718	1-720	720	93

Figure 14

rat_cDNA human_5+3_corrected mus_cDNA_5	CGAGAGACGACAGAAAGTTACGGCTGCGAGAAGACGACAGAGGGTCCAGAAAAGGAAA -----CAGAGGGTCCAGGAAA-GGAAA
rat_cDNA human_5+3_corrected mus_cDNA_5	GTGCTGGAGGGGAGTGGGACAAAAGCAGCGACCAAGTGAATGTCACTTCAGTGACTGAG -----GTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGACATCGCTTCAGTGACTGAA
rat_cDNA human_5+3_corrected mus_cDNA_5	GCCAGGCAAAAAGCGCGGGAAGGATTTGTGTAGCTTGGGACCCCTTCATAGACACTGAT -----GCCAGGCAAAAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTCATAAACACTGAT
rat_cDNA human_5+3_corrected mus_cDNA_5	GACACGTTTACGCAAAATA-GAAATTTGAGGAGAAACGCCTGGGCCTTCGGAAAG---GA -----GACGTGTTTGTGCAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAAGGA
rat_cDNA human_5+3_corrected mus_cDNA_5	GTGATTGATTAGTACTTTCAGTGTAGTTAGGTGACTTTAAGGAGAACTAACTAATGTATACTA -----GTGATCGATTAGTACTTGTAGTTTAGGTGAGTTT--GAGAACTAACTAACTATACTA
rat_cDNA human_5+3_corrected mus_cDNA_5	TTGAGGGAGGAGGAAGAGCATTACAGAGTTTCCAGCAGCAGCAGGAAAGCTTTGGTTAAT -----TTGAGGGAGGAAGGAAGAGCATT-----CCAGCAGCAGCAGGAAAGCTTTGGTTAGT
rat_cDNA human_5+3_corrected mus_cDNA_5	TTGGAAATGGATGATAGCATTAAATAACAGAAGCGCCTCCAGGTCTCTGAAGCTTCAGT -----TTGGAAATGTATGATACCATTAAATAACAGAAGCGCCTCCAGTTCTCTGAAGAGTCAGT
rat_cDNA human_5+3_corrected mus_cDNA_5	CCCCAGCTGAAAGCCAGAAAGACTAAGCCCTAAGCCCTTTTGATCCCTTTGGAAGCA -----CCCCAGCT-----A-GTGTAAGCCTACTAAGCCCTTTTGCTCCCGTTGGAAGCA
rat_cDNA human_5+3_corrected mus_cDNA_5	AAGAACTTTCCTTCCCTGGGTGAAGACTCTCCTCAGAAGATTTCCTGTCTCTGCCTATG -----AAGAACGTTCCCTTCAATCAGGTGAAGGCTCTCCTCAGAAGATTTCCTGTCTCTGCCTATG
rat_cDNA human_5+3_corrected mus_cDNA_5	TTACAAGAGGAATCAAACCAAGACAGAAGAGCTCAGGATGCAGGTGAGAGGCAGGGAG -----ATGAAGGTAAAGGCAGAGGAA TTACAAGAGGATTCAAAGCAAGACAGAAGAGCTCAGGATGCAGAAGAGAGGCAGGGAG *** ** * ***** *
rat_cDNA human_5+3_corrected mus_cDNA_5	TCAGCGGCTTGTGTGCTCTCCCTCACTGCTGTCTGCCTGGTGGTCAOCCCTGGGAGCAGGG TCACTGCTTGTCTGCTCTCCTTTGCTGTGATCTGCCTGGTGGTCAOCCCTGGGAGCAGGG TCAGCTGCTTGTCTGATCTCCTCACTGCCATCTGCCTGGTGGTCAOCCCTGGGAGCAGGG *** * ***** * ***** * ***** * ***** * ***** *
rat_cDNA human_5+3_corrected mus_cDNA_5	CCTGTCTCGCGCTGTGCTGTATGTGCCACAGAGGTGCCTGTACATTTGGGTACC CCTGTCTCGCGCTGTGCTGTATGTGCCACAGAGGTGCCTGTACATTTGGGTACC TCTGTCTCGCGCTGTGCTGTATGTGCCACAGAGGTGCCTGTACATTTGGGTACC ***** * ***** * ***** * ***** * ***** * ***** *
rat_cDNA human_5+3_corrected mus_cDNA_5	TGACCTCCATCCCGATGG-CATCCCGGCCAATGTGGAACGAATAAATTAGGATATAAC TGACTTCCATCCCGACAG-CATCCCGGCCAATGTGGAACGAATAAATTAGGATATAAC TGACCTCCATCCCGAGCGGCATCCCGGCCAATGTGGAACGAATAAATTAGGATATAAC **** * ***** * ***** * ***** * ***** * ***** *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGCCTTACTAGATTGACAGAAAAAGACTTTGATGGCCTGAGCAAACCTGGAGTTACTCATG
AGCTTGGTTAGATTGATGGAAACAGATTTTCTGGCCTGACCAAACCTGGAGTTACTCATG
AGCCTCACTAGATTGACAGAAATGACTTTTCTGGCCTGAGCAGACTGGAGTTACTCATG
*** * ***** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTGCACAGTAATGGCATTTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG
CTTCACAGCAATGGCATTTCACACAATCCCTGACAAGACCTTCTCAGATTTCAGGGCCTTG
CTGCACAGCAATGGCATTTCACAGAGTCAGTGACAAGACCTTCTCGGGCTTGCACTCCTTG
** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGGAGGATACCTTTCTACGGA
CAGGTCTTAAAAATGAGCTATAATAAAGTCCGAAAACCTCAGAAAGATACCTTTTATGGC
CAGGTCTTAAAAATGAGCTATAACAAAGTCCAAATCATTGGAGGATACCTTTGTATGGA
***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCGGGAGCTTGGTCCGCTTGCACTGGATCACAACAACATTGAATTCATCAACCCCTGAG
CTCAGGAGCTTGACACGATTGCACATGGACCAACAATATTGAGTTTATAACCCAGAG
CTCAGGAGCTTGACCCGCTTGCACTGGATCACAACAACATTGAGTTTATCAACCCGAG
*** ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GCTTTTATGGACITACCTCGCTCCGCTTGGTACATTAGAAGGAAACCGGCTCACAAG
GTTTTTATGGGCTCACTTTCTCCGCTGGTGCCTTGGAGGAAATCAGCTCACTAAG
GCGTTTACGGACTCACCTTCTCCGCTTGGTACATCTAGAAGGAAACCGGCTCACAAG
* ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CTCCATCCAGACACATTGTCTCATTAAGCTATCTCCAGATATTTAAACCTCTTTCATT
CTCCACCCAGATACATTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCATT
CTCCATCCAGACACATTGTCTCTTTGAGCTATCTCCAGATATTTAAACCTCTTTCATT
***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGTACCTGTCTGTCTGATAACTTCTT-GACCTCCCTCCCAAAGAAATGGTCTCCTA
AAGTCTCTATACCTGTCTGATAACTTCTT-GACCTCCCTCCCTCAAGAGATGGTCTCCTA
AAGNACCTGTACTGTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTC
*** ** * ***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CATGCCAAACCTAGAAAGCCTGTATTGCAATGGAAACCCATGGACCTGTGACTGCCATTT
TATGCTGACCTAGACAGCCTTTACCTGCAATGGAAACCCATGGACCTGTGATTGCCATTT
TATGCCAAACCTAGAAAGCCTTTACTTGCATGGAAACCCATGGACCTGTGACTGCCATTT
***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAAGTGGTTGTCTGAGTGGATGCAGGGAACCCAGATATAATAAAATGCAAGAAAGACAG
AAAGTGGTTGTCTGACTGGATACAGGNNNNOCAGATGTAATAAAATGCAAAAAGATAG
AAAGTGGTTGTCCGAGTGGATGCAGGGAACCCAGGTA-ACTATCTTGT----TTGTTTG
***** ** *****

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGCTCTTCCAGTCTCAGCAATGTCCCT-TTGCATGAACCCAGGATCTCTAAAGGCA
AAGTCCCTCTAGTCTCAGCAGTGTCCCT-TTGCATGAACCCAGGACTTCTAAAGGCA
TTTCTTTTTTATARKAGTATTTTCTCAATTTCAATTAGATGATATCCCAAAGTC-
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

GACCCCTTGTCTATGTTACCATCTGGAGCTTTCTATGTACAAAGCCAAACATTGATCCAT
AGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCCAGTGTGCCAAGCCAAACATTGACTCAT
-CCCATACCTCCGCCCA-----CTTCCCTACCTACCCATTC-CCATTTTTTGGC
** * * * * ** * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

CAGTGAAGTCAAGAGCCTGGTTACTCAGGAGGACATGGATCTGCTCCACCTCAGCTC
CCCTGAAATCAAGAGCCTGACTATTCTGGAAGACAGTACTTCTGCTTTCATCTCTCCCC
CCTGGCATCCCC-----
* * * * *

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AAGATTCATAGAAACCTTTGGCTCCTTGTCTTTGAACATGACANANNNTCTGGAATA
AAGTTTCATAGGCAACCTTTGGCTCCTCAGTTTGAATATGACAGATCAGTCTGGAATG

rat_cDNA
human_5+3_corrected
mus_cDNA_5

AGGCCGACATGGTCTGTAGTATCCAAAGCCATCAAGGACATCAACCACTGCATTCACTG
AAGCTAACATGGTCTGCAGTATTCAAAGCCCTCAAGGACATCAACCACTGCATTCACTG

AAGAAAATGACTACATCATGCTAAATGCGTCATTTTCCACAAATCTTGTGTGCAGTGTAG
AAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTTTGGTGTGCAACATAG

ATTATAATCACATCCAGCCAGTGTGGCAACTTCTGGCTTTATACAGTGA CTCTCTCTGA
ATTACGGTCACATTCCAGCCAGTGTGGCAAAATTTTGGCTTTGTACAGTGATTCTCTCTCTGA

TACTAGAAAGGAAGCCCCAGCTTACCGAGACTCCTTCACTGTCTTCTAGATATAAACAGG
TACTAGAAAGGAGCCACTTGCTTAGTGAAACAACCGCAGCTCTATTACAATATAAACAGG

TGGCTCTTAGGCTGAAGACATTTTACCAGCATAGAGGCTGATGTCAGAGCAGACCOCTT
TGGCTCCTAAGCTGAAGACATTTTACCAACATAGAGGCAGATCTCAGAGCAGATCOCT

TTTGGTTCACACAAGAAAAATTTGTTCTTGCACTGAAACAGAACTGCCACCACACTTAGCA
CTTGGTTAATGCAAGACCAAAATTTCTTTCAGCTGAAACAGAACTGCCACCACATTAGTA

CATTACAGATCCAGTTTTCCACTGATGOTCAAATCGCTTTACCAAGGGCGGAGATGAGAG
CATTACAGATCCAGTACTCCAGTGTGOTCAAATCACTTTACCAAGAGCAGAGATGAGGC

CGGAGAGACTCAAATGGACCATGATCCTGATGATGAACATOCCTAACTGGAAOGCCTG
CAGTGAAACACAAATGGACTATGATTTCAAGGGATAACAATACTAAGCTGGAACTACTG

TCTTGGTTGGCGGCACCTATTGCCCTGAGCTGTCCAGGCCAAGGGGADCCCTTCACTCACT
TCTTGGTAGGTGGACCGTTGGCCCTGAAGCTGCCAGGCCAAGGAGACCCCAACCCACAGG

TGGAAATGGCTTCTAGCTGATGGGAGTAAAGTGAGAGCCCTTAAGTTAGGAGGATGGGC
TGGATTGGCTTCTAGCTGATGGAAGTAAAGTGAGAGCCCTTATGTCACTGAGGATGGAC

GAATCCTAATAGACAAAATGGGAAGTTGGAACTGCAGATGGCTGACAGCTTTGATGCAG
GGATCCTAATAGACAAAAGTGGAAAATTGGAACTCCAGATGGCTGATAGTTTTCACACAG

GTCTTTACCACTGCATAAGCAACATGATGCAGATGCGGATGTTCTCACATACAGGATAA
GCGTATATCACTGTATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAA

CTGTGGTAGAGCCCTATGGAGAAAGCACACATGACAGTGGAGTCCAGCACACAGTGGTTA
CTGTGGTAGAAOCTTTGGTCGAAGCCTATCAGGAAAATGGGATTCATCACACAGTTTCA

CGGGTGAGAGGCTGGACCTTGCATGCCCTTCCACGGGTGTTCCAGATGCTTCTATTAGCT
TTGGTGAAACACTTGATCTTGCATGCCCTTCTACTGGTATCCAGATGCCCTATTAGCT

GGATTCTTCCAGGGAACTGTGTTCTCTCAGCCATCAAGAGACAGGCAAATTCCTAAC
GGGTTATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCCTAAC

rat_cdna human 5+3 corrected mus_cdna_5	ATGGGACCTTAAGAATATTACAGGTTACGCCAAAAGATCAAGGTCATTACCAATGTGTGG ATGGCACATTAAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTGG
rat_cdna human 5+3 corrected mus_cdna_5	CTGCCAACCCATCAGGGGCCGACTTTTCCAGTTTTAAAGTTTCAGTTCAAAAGAAAGGC CAGCCAACCCATCAGGGGTTGATTTTTGATTTTCCAAGTTTCAGTCAAGATGAAAGGAC
rat_cdna human 5+3 corrected mus_cdna_5	AAAGGATGGTTGAGCATGACAGGGAGGCAGGTGGATCTGGACTTGGAGAACCCAACTCCA AAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGATCCAACTCTA
rat_cdna human 5+3 corrected mus_cdna_5	GTGTTTCCTTAAGCAGCCAGCATCTTTGAACTCTCTGCATCAGCTTTGACAGGTCAG TTGCTCATCTTAAGGAGCCACAGGTGCACAACCTCCGTACATCTGCTCTGATGGAGGCTG
rat_cdna human 5+3 corrected mus_cdna_5	AGGCTGGAAACAAGTCTCCGCTGTACATAGGAAGAACAAACATAGAGACTTAATACATC AGGTTGGAAACACACCTCAAGCACAGTAAGAGGCACAACCTATCGGGAATTAACACTCC
rat_cdna human 5+3 corrected mus_cdna_5	GGCGGCTGGGGATTCCACGCTCCGGGCAITCAGGGAGCATAGGAGGCAGCTCCCTCTCT AGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAAATAGGAGGCATTTCCCTCCCT
rat_cdna human 5+3 corrected mus_cdna_5	CTGCTCGGAGAATTGACCCGCAACGCTGGGCGACACTTCTAGAAAAGCCAAAAGAAAT CTGCTAGGAGAATTGACCCACAACTTGGGCGGCACTGTTGGAGAAAGCTAAAAAGAAATG
rat_cdna human 5+3 corrected mus_cdna_5	CTGTGCCAAAAAGCAAGAAATACCAAGTAAGCCAGTGCCACTGGCTGTTCCTCTG CTATGCCAGACAGCGAGAAATACCAAGTGAGCCCGACCCCGAGTGGTCAACCCACTCC
rat_cdna human 5+3 corrected mus_cdna_5	TGGAACTCACTGACGAGGAAAGGATGCCCTCTGGCATGATTCCTCCAGATGAAGAATTCA CAACATACCTGGTGAAGAAGACGATTCTCAGGCATGCTCGCTCTACATGAGGAATTGA
rat_cdna human 5+3 corrected mus_cdna_5	TGTTCTGAAAACCTAAGGCTTCTGGTGTCCAGGAAGGTCACCAACTGCTGACTCTGGAC TGTTCCCGCCACTAAAGCTTTGAACCTTCCAGCAAGGACAGTGACTGCTGACTCCAGAA
rat_cdna human 5+3 corrected mus_cdna_5	CAGTAAATCATGGTTTTATGACGAGTATAGCTTCTGGCACAGAAGTCTCAACTGTGAATC CAATATCTGATAGTCCTATGACAAACATAAATTATGGCACAGAAGTCTCCGTTGTGAATT
rat_cdna human 5+3 corrected mus_cdna_5	CACAAACACTACAATCTGAGCACCTTCTGATTCAAAATATTTAGTGAACAAAGGTA CACAAATACTACCACCTGAAGAACCACAGATTCAAACTGTCTACTGCTATTAAACTA
rat_cdna human 5+3 corrected mus_cdna_5	CAGCTGTGACAAAGAGTATGAACCCATCCATAGCAAGCAAAATAGAAGATACAACCAACC CAGCCATGTCAAAGAATATAAACCACCAATGTCAAGCCAAATACAGGCACCAACCAATC
rat_cdna human 5+3 corrected mus_cdna_5	AAAACCCAATCATTTATCTTTCCATC-----AGTAGCTGAATTCGAGATTCTG-CT AACATTCATCCACTGTCTTTCCACTGCTACTTGGAGCAACTGAATTTCAGGACTCTGAC

CAGGCA---GGAAGAGCATCTTCCAAAGTGCAACCCCTGTAACAGGGGGAAACATGGCT
 GAGGGAAGAGGAAGAGAGCATTTCC--AGTA-ACCCCCATAACAGTAAGGACTATGATC

ACCTATGGCCATACCAACACATA--TAGTAGCTTTACCAGCAAAGCCAGTACAGTCTTGC
AAAGATGNTCAATGTCAAANATGCTTAGTAGCAACCACCAACAA--CTATTA-----TTAG

AGCCAATAAATCCACGAAAGTTATGGA~~CT~~CAGATACCTATTACAGGAGTCAGCAGAC
AGTCAGTAAATACCAAAATAGTCAT-----CAGACATCTGTAAGAGAGAGTGAAC

CTAGCAGTAGTGACATCTCTTCTCACACTACTGCAGACCTAGCTTCTCCAGTCACCCTT
CCAGGCACCAATCACTTCTATTCTCACACTACTCAAATACCTAGCACCTGCAGCTTCCCTT

CAGGTTACACACCACTGCGTCGTCCTTATTTACATTCCCTAGAAACACCAATACAGGTA
CAGATCCACACACAGCTGCTCAITCTCAGTTTCOGATCCCTAGANNNAATAGTACAGTTA

ACTTCCCCCTTGTCAGGCCTTGGGAAGACASAGGACAATTGGAGCAGAGGGAGAGTTA
ACATCCCGCTGTTCAAGCCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGGGGATTA

AAAACCCACATAGAAACCCAGTTCTCCGACGGCATAGACACAGGACTGTGAGGCCAGCAA
TCAGCCCATATAGAACTCCAGTTCTGCCAGGGCATAGATACAGCATTTCAGGTCAACAA

TCAGGGACCTGCTAACAAAAATGTGAGCCAAAGTTCCAGCCACAGAGTAOCTGGGATGT
CCAGAGGTTCTCTGAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACAT

GCCACACATGTCCITCCGCAAGGGGGTCCACAGTGGCTACTGCAGCACTGTCAATTCCAA
GTCTGTCTCTGTCTTCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAA

GTTCAATCCACAGTGGCCCTCCCCAAACTAATAATGTTGGGGTCATAGCAGAAGAGTCTA
 GTGCTGCTCCCATCACCTTCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCTA

CCACTGTGGTCAAGAAACCACTGTTACTATTTAAGGACAAACAAAATGTAGATATTGAGA
CAACTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGANNNGAAA

TAATAACAACCACTACAATAATATTCGGGAGGGGAAAGTAACCAAGTGATTCCTACGGAAG
NNRCAACACCCACAATAAAATATTCAGGACTNGAAATTTCCCAAGTGACTCCAACCTGGTG

CAAGCATGACTTCTGCTCCACATCTGTATCOCTGGGGAAATCTCCTGTAGACAATAGTG
CAGTCATGACATATGCTCCAAATCCATACCGATGGAAAACTCACAAAGTAAAGOCA

GTCACCTGAGCATGCCTGGGACCATCCAACTGGGAAAGATTCAGTGGAAACATACCCAC
GTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTCAGTGATTACATCGTCAC

rat_cDNA human_5+3 corrected mus_cDNA_5	TTCCGAGCCCCCTCAGCACACCCCTCAATAACAA-----CAAGCACAAAATTTCTCAAAGA TTTCAGGTGCTATCAACCAAGCCCAATGACTATTATAGOCATTACAGGTTTTCAAGAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GGAAAACTCCCTTGCCACAGATCTTTGTAAATAACCAAGAAGSAGGGGATGTTAAAGA GGAAATTCCTGGCAACAGAACTTTGTAAATAACCATACCCAAAAGGCAGATTAAAGA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ATCCATATCAATTGGTTTACAAAAGAACCCAGCCGCAAGOTTCCCAAAATAGCTCCTC ATCAACATAAAGTTAGTTTACAAAAGGCACAGCTGTGATGCTTCCTAAACATCTCCTG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TTTACCCACAG-GTCAGAGTTCCCCCTCAGATTCTACAACCTCTTTGACAAGTCCGCCA CTTTACC-ACAGAGACAAAGTTCCCTTTCCATTTCAACCACTTTCAACAAGTGTGATG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CCAGCTCTGTCTACAACAATGGCTGCCACTCAGAACAAGGGCACTGAAGTAGTATCAGGT CAAAATTCATCTAATACCTTGACTACCGCTCACCACTACGACCAAA--ACACACAA-T -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GCCAGAAGTCTCTCAGCAGGGAAGAGCAGCCCTTCAACAACTCCTCTCCAG-TGCTTCC CCTGGAAGTCTTCCAACAAGAAGGAGCTTCCCTTC-CCACCCCTTAACCTATGCTTCC -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TAGCOCATAAGCAAGAGATCTAATACATTAACTTCTTGTCAACGGAAACCCCCCAGT TAGTATTATAAGCAAGAGACTCAAGTACAAAAGCATCATATCAACGCAACAGCAACCGC -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GACAAGTCTACTGCTACTGCTATCTGTCTATTATGTCTGAAACCAACGAAACAGATCCAA AACAACTCTACCTTCCCTGCTATCTGTCTACTTATGAACCAACAGAGAGATCTAG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	AGAAGCAAAAGACCAATATAAGGGGCTCG---GAAGAACAAGAACGCAACACCCAC AGCACAACAATAACAAGAGAAGGAOCTCAAAAGAGAAGCAGGACTGACCCAAACATCTC -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CCCCAGGCAGGTTTCTGGCTATAGTGCACTACTCAGCTCTAACACAGCTGATAACCCCTT TCCAGACCAAGATTCTGGCTTCACTACACCACTGCTATGACNACCTCCTMNGCTCTNNN -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GGCTTTCAGTCAATCCCCACGACAAGATGATGGTGGAAATGTAAGTGCAAGTTGCTTATCA NGCATTCACTCATTCOCCACCAGAAAACACAACCTGGGATTTCAAGCACATCAGTTTTTCA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CTCAACAACCTCTCTTCTGGCCA---TAACTGAAGTGTGAG-AAGTAC--ACCCAGAC TTCAAGAACTCTTAATCTGACAGATGTGATTGAGAAGTACCCCAAGCAGTACTCAGAC -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TTTGGGAATACACAGCTTTGGAACAACGTTGTTGAGCAATCACAGGAGAGTACCAC TTTGAAGAGCACAAATGCTTCTGAAACAACCTTTGTCAGCAATCACACCAGAGTACCAC -----
rat_cDNA human_5+3 corrected mus_cDNA_5	AGTGAAGAGAGCTCA---GACAC---ACCAACCACTCCTCAGCAGTGGGGGGCCCCC AACTAGGAAGCATCATTAGACACTCAACCAACCACTTCTTGAAGCAGTGTCTACTCT -----

U
T
E
R
I
N
E
R
Y
C
O
D
E
S

rat_cDNA human 5+3 corrected mus_cDNA_5	AGTGCCCACTCCTTCCCCACCTCCTTTTACTAAGGGTGTGGTTACAGACAGCAAGTCAC AATGCCAGTTCOCATCTCCCTCCCTTTACTCAGAGAGCAGTTACTGACAACTGGGAC
rat_cDNA human 5+3 corrected mus_cDNA_5	ATCAGCTTTCAGATGACGTCAAATAGAGTGGTCACCATATATGAATCTTCAAGGCACAA TOCCATTTCCGGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAA
rat_cDNA human 5+3 corrected mus_cDNA_5	TACAGATCTGCAGCAACCTCAGCAGAGGCTAGCCCCAATCCTGAGATCATACTGGAAC TCNNNNNNNCA--AATGCCAAGTTCA-----CNAATTGNGAACNNNNNACTCNNNN
rat_cDNA human 5+3 corrected mus_cDNA_5	CACTGACTCTCCCTCTAATCTGTTCCATCCACTTCTGTGCCAGCACTAAGGGTAGATAA NACTTCATCTACNTCTAATCTGTTACATTCTACTCCCATGCCAGCACTAACCAAGTTAA
rat_cDNA human 5+3 corrected mus_cDNA_5	ACCACAGAATTCTAAATGGAAGCCCTCTCCCTGGCCAGACACAAATATCAGCTCAAGTC ATCACAGAATTCCAAATTAAGTCCATCTCCCTGGGCGAATAACCAATTTGGCACAACCC
rat_cDNA human 5+3 corrected mus_cDNA_5	ATACTCCGAAACCATTTAGAGAGGGCAAAAGGCCAGCAGTAAGCATGTCCCCCACTCAG ATACTCAGACATTGCTGAAAAGGCCAAAAGCCAGAAGTAAGCATGTTGGCT-ACTACAG
rat_cDNA human 5+3 corrected mus_cDNA_5	-CCTTCCAGAGGCCAGCACTCATGCCCTCACACTGGAATACACAGAAGCATGCAGAAAGA GCCGTGTCGAGGCCACCCTCTTGTTCAGATTGGGATGGACAGAGAACACAAAGAGA
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGTTTTTGATAAGAAACCTGGTCAA-AACC--CAACTTCCAAACATCTGCCCTAAGTCT GTGACTTTGATAAGAAACAGTTCAGAAGCAACCACTTCCAACTCCTTCCCTTTGACT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTCTACCTAAGACTCTATTGAAAAGCCAGAAATAATTGGAGGAAAGGCTGCAAGCTTTA CTTTGTCTAGGTATATATTGAAAAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTAA
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGTTCAGCTAATTCAAGCTTTTCTTCTCTGAGGGCTGTTGGAGACCACTGCCCA CTATTCCAGCTAACTCAGATGCCCTTCTTCCCTGTGAAGCTGTGGAAATCCCTGCCCA
rat_cDNA human 5+3 corrected mus_cDNA_5	TCATCCACTGGACAGAGTTTCATCAGGANTTGAATATCCCAAGGGACACAGAAAAGCC CCATTTCATTGGACAGAGTNNNNTCAGGACTTGATTATCTAAGAGGAAACAGAAATAGCA
rat_cDNA human 5+3 corrected mus_cDNA_5	GGTTCACGTCCTTCCCAATGGCACCTTGTCCATCCAGAGGGTCAGTATTCAGGACCGTG GGGTCCAGGTTCTCCCCAATGGTACCTGTCCATCCAGAGGGTGGAATTCAGGACCGCG
rat_cDNA human 5+3 corrected mus_cDNA_5	GACAGTACCTGTGCTCTGCATTTAATCCACTGGGCGTAGACCATTTTCATGCTCTTTGT GACAGTACTTGTTGTTCCGCATCCCAATCTGTTTGGCAGACACCACTTCATGTCACCTTGT
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTGGTTTTTTAOCOGGCAAGGATTTTGGACAGACATGTCAAGGAGATCAGTTTCACT CTGTGGTTTCTATCTCTCCAGGATCCTGGAGAGACGTACCAAGAGATCAGTTTCACT

rat_cDNA human_5+3 corrected mus_cDNA_5	TTGGAAGTACTGTGGAACAAAGTGCAGAGTGGAGGGTATGCCGAGGCCTACGGTTTCCT CCGGAAGCACTGTGGAACGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTAOCCT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GGATACTTGCAAACCAAACGGTGGTCTCAGAAACGGCCAGGGAAGCAGAAAGGTCTGGG GGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCAGGGAAGTAGGCAGGCTGTGG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TAAACCTGATGGAACATTGATCATCTATAATCTGAGTCTTTATGATCGTGGTTTTTACA TGACGGTTGACGGAACATTGGTCCCTCCCAATCTCAGTATTTATGACCGTGGCTTTTACA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	AGTGTGTGGCCAGCAACCCATCTGGCCAGGATTCAGTGTGGTTAAGATACAAGTCATCA AATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCAGTGTGGTTAAATACAGTCATTG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CAGCTCCOCCTGTCTATTATAGAGCAAAAGAGGCCAAGCCATCGTTGGGGTTTTAGGTGGAA CAGCACCACCTGTTATTCTAGAGCAAAAGAGGCCAAGTCATTGTAGGCACCTTGGGGTGAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GTTTGAAACTGCCCTGCACTGCAAAAGGAACCTCCCGCCCTAGTGTTCAGTGGGTCTTT GTTTAAACTGCCCTGTACTGCAAAAGGAACCTCCTCAGCCAGCGTTTACTGGGTCTCT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ATGATGGGACTGAACATAAAACCATTCAGTTGACTCATTCAGATTTTTCTGTATCCAA CTGATGGCACTGAAGTGAAACCATTCAGTTTACCAATTCAGTTGTTCTTATTTTCAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ATGGAACCTCTGTATATAGAAGCATCGCTCCTTCAGTGGGGGCACCTTATGAGTGCATTG ATGGGACTTTGTATATAGAAGCATCGCTCCTTCAGACAGGGGCACCTTATGAATGCATTG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CCACAGCTCCTCAGGCTCAGAGAGAAGGGTAGTGATTCTTACTGTGGAAGAGGGAGAGA CTACAGTTCCACTGGTTTCGGAGCGAAGAGTAGTAATGCTTACAATGGAAGAGCGAGTGA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CAATCCCCAGGATAGAACTGCTCTCAGAAATGGACTGAGGTGAATTTGGGTGAGAAAT CCAGCCCCAGGATAGAGCTGCATCCAGAAAGGACTGAAGTGAATTTGGGGACAAAT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TACTACTGAACTGCTCAGCTACTGGGGATCCAAAGCCTAGAATAATCTGGAGGCTGCCAT TACTACTGAACTGCTCAGCCACTGGGGAGGCCAAACCCCAATAATGTGGAGGTACCAT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	CCAGGCTGTCTACGACCACTGGCCACAGAATGGGCAGCCGAATCCAGCTACCCAAATG CCAAGGCTGTGGTCGACCACTGG-----GCAGCTGGATCCAGTCTACCCAAATG -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GATCCTGTGGTGGTGGTTCAGTGACGGAAAGACGCTGGTGACTACTTATGTGTGGCAA GATCCTGTGTATTTGGATCAGTAACAGAAAGACAGTGGTGTCTACTTGTGTGTGGCAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GAAACAAATGGGAGATGACCTAGTCTCTGATGCATGTCCGCCTGAGATTGACACCTGCCA GAAACAAATGGGGGATGATCTGATACTGATGCATGTAGCCTAAGACTGAAACCTGCCA -----

rat_cDNA human 5+3 corrected mus_cDNA_5	AAATTGAACAGAAAGCAGTATTTTAAAGCAAGTGCTCCATGGGAAAGATTTCAGTTG AAATTGACCCACAGCAGTATTTTAGAAGCAAGTGCTCCATGGGAAAGATTTCAGTTG
rat_cDNA human 5+3 corrected mus_cDNA_5	ACTGCAAGGCTCTGGCTCCCTGTGCTGAGGTATCTGGAGTTTGCTGATGGGACAG ATTGCAAGCTTCCGGCTCCCTAGTGCCAGAGATATCTTGGAGTTTGCTGATGGAAOCA
rat_cDNA human 5+3 corrected mus_cDNA_5	TGCTCAACAATGTAGCCCAAGCTGATGACAGTGGCTATAGGACCAAGAGGTACACCTTT TGATCAACAATGCAATGCAAGCCGATGACAGTGGCCACAGGACTAGGAGATATACCTTT
rat_cDNA human 5+3 corrected mus_cDNA_5	TCCACAATGGAACTTTGATTTTCAACAAGCTTGGGATGGCAGAGGAAGGAGATTATATCT TCAACAATGGAACTTTTATCTTCAACAAGTTGGGATAGCGGAGGAAGGAGATTATATCT
rat_cDNA human 5+3 corrected mus_cDNA_5	GCTCTGCCAGAACACCTTAGGGAAGATGAGATGAAGTCCACCTAACAGTTCTAACAG GCTATGCCAGAACACCTTAGGGAAGATGAAATGAAGTCCACTTAACAGTTATAACAG
rat_cDNA human 5+3 corrected mus_cDNA_5	CCATCCACGGATAAGGCAAGCTACAGACCACCATGAGGCTCAGGGCTGGGAAACAG CTGCTCCCGGATAAGGCAGAGTAACAAACCAACAGAGAAATCAAAGCTGGAGACACAG
rat_cDNA human 5+3 corrected mus_cDNA_5	CTGTCTTGACTGCGAGGTCACTGGGGAACCGAAGCCAAATGTATTTGGTTGCTGCCTT CTGTCTTGACTGTGAGGTCACTGGGATCCCAACCAAAATATTTGGTTGCTGCCTT
rat_cDNA human 5+3 corrected mus_cDNA_5	CCAACAATGTCATTTTCTTCTCCATGACAGGTTCACTTTTATGCCAATAGACTTTGT CCAATGACATGATTTCTTCTCCATGATAGGTACACATTTATGCCAATGGGTCTTTGA
rat_cDNA human 5+3 corrected mus_cDNA_5	CCATCCATAAAGTGAAACCACTTGACTCTGGGACTATGTGTGCTAGCTCAGAACTTA CCATCAACAAAGTGAAACTGCTCGATTCTGGAGAGTAAGTATGTGTAGCCCGAAATCCA
rat_cDNA human 5+3 corrected mus_cDNA_5	GTGGGATGACACTAAGACATACAACTGGACATTGTCTTAAACCTCCATTATCAATG GTGGGATGACACCAAAATGTACAACTGGATGTGTCTTAAACCTCCATTATCAATG
rat_cDNA human 5+3 corrected mus_cDNA_5	GCCTGTATGCAACAAGACTGTTATTAAGCCACAGCCATTGGGCACTCCAAAAATACT GTCTGTATACAAACAGAACTGTTATTAAGCCACAGCTGTGAGACATTCCAAAAACACT
rat_cDNA human 5+3 corrected mus_cDNA_5	TTGACTGCAGAGCAGATGGGATCCCATCTTCCAGGTACGTTGATTATGCCAGGCAATA TTGACTGCAGAGCTGAAGGGACACCATCTCTGAAGTATGTGGATCATGCCAGCAATA
rat_cDNA human 5+3 corrected mus_cDNA_5	TTTTCTCCAGCTCCATACTTTGGAGCAGAGTCACGGTCCATCCAAATGAACTTGG TTTTCTCAGAGCCCATACTATGGAAGCAGAAATCAGTCCATAAAATGAACTTGG
rat_cDNA human 5+3 corrected mus_cDNA_5	AGATGAGGAACATCCGGCTTCTGACTCTGCGGACTCACTGTGTGGTTCGGAGCGAGG AAATTAGGAATGTGAGGCTTTCAGATTCAGCGACTTTATCTGTGTGGCCGAAATGAAG

rat_cDNA human 5+3 corrected mus_cDNA_5	GAGGAGAGAGTGTGTTGGTAGTGCAGTTAGAAGTCTAGAAATGCTGAGAAGACCAACAT GTGGAGAGAGCGTGTGGTAGTACAGTTAGAAGTACTGGAAATGCTGAGAAGACCGACAT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCAGAAACCCATTCAACGAAAAAGTCATCGCCCAAGCTGGCAAGCCCGTAGCACTGAAC TTAGAAATCCATTTAATGAAAAAATAGTTGCCAGCTGGGAAAGTCCACAGCATTGAATT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	GCTCTGTGGATGGGAACCCCACTGAAATTAOCTGGATCTTACCTGAOGGCACACAGT GCTCTGTGATGGTAACCCACCACTGAAATAATCTGGATTTTACCAATGGCACACGAT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TTGCTAACAGACCAACCAATTOCCGATCTGATGGCAGGCAATGGCTCTCTCATCTTT TTTCCAAATGGACCAAAAGTTATCAGTATCTGATAGCAAGCAATGGTTCTTTTATCATT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACAAAGCAACTCGGAACAAGTCAGGGAAGTATCGCTGTGCAGCCAGGAATAAGGTTGGCT CTAAACAACCTCGGGAGGATGCAGGAAAATATCGCTGTGCAGCTAGGAATAAGTTGGCT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACATCGAGAACTCATCTGTAGAGATTGGGCAGAAGCCAGTCATTCTGACATAAGAAC ATATTGAGAAATTAGTCATATTAGAAATGGCCAGAAGCCAGTTATTCTTACCTATGCAC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	CAGGGATGGTGAAGAGCGTCAGTGGGGAACCGTTATCACTGCATTGTGTCTGATGGGA CAGGGACAGTAAAAGGCATCAGTGGAGAATCTCTATCACTGCATTGTGTCTGATGGAA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCCCCAAGCCAAATGTCAAGTGGACTACACCGGGTGGCCATGTAATCGACAGGCGCTCAAG TCCCTAAGCCAAATATCAATGGACTATGCCAAGTGGTTATGTAGTAGACAGGCGCTCAA -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGGATGGAAAATACATACTGCATGAAATGGCAGCTGGTCATCAAAGCAACACAGCTC TTAATGGGAATACATATTGCATGACAAATGGCACTTAGTCATTAAAGAGCAACAGCTT -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ACGACCAAGGAAATATATCTGTAGGGCTCAAAACAGTGTGGCCAGGCAGTTATTAGCG ATGACAGAGGAACTATATCTGTAGGGCTCAAATAGTGTGGTCATACACTGATTACTG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TGTCAGTGATGGTTGTGGCTACCGTCCCGAATCATAACTAOCCTACCCAGGAACATGC TTCCAGTAATGATTGTAGCTACCGTCCCGAATTACAAATGTCCACCCAGGAGTATTG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	TCAGGAGGACAGGGGAAGCCATGCAGCTCCACTGTGTGGCTTGGGAATCCCCAAGCCAA TCACCAGGACAGGGGCGCCCTTTCAGCTCCACTGTGTGGCTTGGGAGTTCCCAAGCCAG -----
rat_cDNA human 5+3 corrected mus_cDNA_5	AAGTCACCTGGGAGCGCCAGACACTCCCTGCTCTCAAAGCAACAGCAGAAACCCC AAATCACATGGGAGATGCCTGACCACTCCCTCTCTCAACGGCAAGTAAGAGAGGACAC -----
rat_cDNA human 5+3 corrected mus_cDNA_5	ATAGAAGTGAGATGCTTACCCACAAGGTACGCTGGTCATTGAGAACTCCAAACCTCGG ATGGAAGTGAGCAGCTTCACTTACAGGTACCGTAGTCATTGAGAACTCCAAACCTCGG -----

rat_cDNA human_5+3 corrected mus_cDNA_5	ATTCCGGAGTCTATAAGTGCAGAGCTCAGAACCTACTGGGACTGATTACGCAACAACCTT ATTCTGGGATATACAAATGCACAGCAAGAAOCCACTTGGTAGTGATTATGCAGCAACGT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ACATCCAGGTACTCTGACAGGAGGGGAGACTAAAATTCAACAGAACTCCACATCCACA ATATTCAAGTAATCTGACATGAA-----ATAATAAAGT-CAACAA-----CATCTGGGCA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GGGTTTATTTTTTGAAGAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAGTCTGA GAATTTATTTTTTGAAGAAGTTTAATCAAAGGCAGCCATAGGCATGTAAATGAATTTGA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ATACATTACAGTATTAAATTTACAATGGACATGCGA--TGA---GACTTGTAATGAAA ATACATTACAGTATTAAATTTACAATGAACATGCAAAATAAAAGGACTTGTAATTAAT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GCATTGTGAAC TGAA---ACCGAGTCTCTG--TGGATCTCAAAGCAAACCTCTTAACCTAA GCATTATGAAC TGATGACTGATTTTATTTAATGGATCTCAAACAAACCTTTAACTTAA -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GGCACTTTGATTTTTGCCAACAAATAATAACAAACATTAAGAGAAAAAATGATCCACTAC GGCACTTTTATTTTTGCCAACAAATAACATAACAA-----ACATTGAAACGGTTCACTAT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	GAAATAACAAACGGCTAATGCACCTGAATTCT-CAGTAAAAAGACCTTTCTCTCGCTAAC AAAATAACAAATGGCTAATGTACCTGAATTTTTTCACTAAAAA---TGAACCTT-CTAAT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	AGTTGCCAGCTGCCTCGTGTCTGTTTCTACCAATGTCAAAACATCGCACACAGGGTGA A---CCAGTTGCCTAGTGTCCACCTCTATCAATGTTACAAGCATGGCACTCAG----A -----
rat_cDNA human_5+3 corrected mus_cDNA_5	ATGGAGTCAACGGGAAGATTAGTTTGGGCTGTGTAAATCTCAATGTACAAATATTC ACAGAGCAATGGAAATATTAAATCTGC-----AATCT--ATGTATAAATATTT -----
rat_cDNA human_5+3 corrected mus_cDNA_5	TGTCNCTGGTTTATAAATTTTT-GATAAAACGGAAAAAATAAAAAAAAAAAAAA TGT----GGTTTATAAATTTTTTGCTAAACCTACAGAAATAG----- -----
rat_cDNA human_5+3 corrected mus_cDNA_5	AAAA ----- -----

(rat_cDNA: SEQ ID NO:7)
(human_5+3 corrected: SEQ ID NO:8)
(mus_cDNA_5: SEQ ID NO:9)

rat
human_5+3_corrected
mouse_5_corrected

MQVRGREVSGLLISLTAVCLVVTGPSRACPRRCACYPTVEVHCTFRYLTSIPDGI PANVE
MKVKGRGITCLLVSEFAVICLVATPGGKACPRRCACYMPTVEVHCTFRYLTSIPDSIPP NVE
MQKRGREVSCLLISLTAICLVVTGPSRVCPRRCACYPTEVHCTFRDLTSPD--GPANVE

rat
human_5+3_corrected
mouse_5_corrected

RINLGYNLSLTRITENDFDGLSKLEILMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
RINLGYNLSLRLMETDFSGLTKELEIMLHSNGIHTIPDKTFSSDLOALQVLKMSYNKVRKL
RVNLGYNSLTRITENDFSGLSKLEILMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII

rat
human_5+3_corrected
mouse_5_corrected

RKDFTFYGLGSLVRLHLHDENNIEFINFEAFYGLTSLRLVHLEGNRLTKLHPDTFVSLSYLQ
QKDFTFYGLRSLTRLHMDHNENIEFINFEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYLQ
EKDFTLYGLRSLTRLHLHDENNIEFINFEAFYGLTTLRLVHLEGNRLTKLHPDTFVSLSYLQ

rat
human_5+3_corrected
mouse_5_corrected

IFKTSFIKYLEFLSUNFLTSLPKEMVSYMPLNLSLYLHGNEPWTCDCHLNWLSSEWMQGNPD
IFKISFIKFLYLSINFILTSLPKEMVSYMPLDOSLYLHGNEPWTCDCHLNWLSDWIQ--PDV
IFKTSFIKXLYLYDNF-TSLPKEMVSSMPNLSLYLHGNEPWTCDCHLNWLSSEWMQGNP--

rat
human_5+3_corrected
mouse_5_corrected

IKCKKDRSSSSPOQCPLQNPRI SKGRPFFAMVP SGAF LCTKPTIDPSLKSKSLVTOEDNG
IKCKKDRSPSSAQQCPLQNPRT SKGKPLAMVSAAPFOCAKPTIDS SLKSKSLTILEDSS

rat
human_5+3_corrected
mouse_5_corrected

SASTSPQDFIEFPFGSLNMXTKXSGNKADMVCS IQKPSRTSPTA TEENDYIMLNASFST
SAFISPQGFMAPPFGSLTLMNTDQSGNEANMVCSI QKPSRTSPTA TEENDYIVLNTSFST

rat
human_5+3_corrected
mouse_5_corrected

NLVCSDVDYNIHQPVWQLALYSDSPLILERSKPQTETPSLSSRYKQVALRPEDIFTSLIA
FLVCNIDYGHIOQPVWQIALYSDSPLILERSHLLSETPOLYYKYKQVAPKPEDIFTNIEA

rat
human_5+3_corrected
mouse_5_corrected

DVRADFFWFQQEKTIVLQINRTATTTLSTLQIQFSTDQAIALFRAEMRAERLKWMTMI LMNN
DLRADPSWLMOQDISLQINRTATTFTSLQIQYSSDAQITLFRAEMRPVKHKWTMISRDN

rat
human_5+3_corrected
mouse_5_corrected

PKLERTVLVGSTIALSCPGRGDPS PHLENLIADGSKVRAPTVSEDGRIL IDKNKGKLELOM
TKLEHTVLVGTVGLNCPCGGDPTPHVIDWILLADGSKVRAPTVSEDGRIL IDKSGKLELOM

rat
human_5+3_corrected
mouse_5_corrected

ADSFDAGLYHCISTNDADADVITYRITVVEPYGESTHDSGVQHTVVTTGETLDLPCLSTGV
ADSFDTGvyhcissnyddadilTYRITVVEPLVEAYOENGIIHTVFIGETLDLPCHSTGI

rat
human_5+3_corrected
mouse_5_corrected

PDAISWILPGNTVFSQP SRDRQIILNNGTLRILQVTPKOOGHYQCVAA NPBGADFESSFKV
PDAISWVPIPGNNVLYQSSRDKKVLNNGTLRILQVTPKDGYYTRCVAA NPBGVDLIFVQ

rat
human_5+3_corrected
mouse_5_corrected

SVQKKGQRNVHEHREAGSGGLGE PNSSVSLKQPASLKL SASALTGSEAGKVQSVGVRKKN
SVKMKGORPLEHDEGETEGSGLDES NPIAHLEKPPGAQLRTSALMEAEVGKHTSSTSKRN

rat
human_5+3_corrected
mouse_5_corrected

HARDLIHRRRGDSTLRFRFHRRQLPLSARRIDPQRWAALLEKAKKN6VPKKQENTTVKPV
YRELTLQRRGDSTHRRFRNRHFP SARRIDPQHWAALLEKAKKNAMPDKRENTTVSPF


```

rat      WPEHKYQLKSYSETIEKGRPAVSMSPHLSPLEASTHSHWNTQKPAKSVFDKKPGQNF
human_5+3_corrected WAAYQIWEHKPYSDIAEKGGKPEVSMLATTLGLSEATTLVSDWDGGRKNTKKSDFDKKPVQEA
mouse_5_corrected  -----

```

rat	-TSKHLPYVSLPKTLLKKPRIIGGKAASFTVPANSDVFLPCEAVGDPLPIIHWTRVSSGX
human_5+3_corrected	TTSKLLPFDLSRYIFEKPRIVGGKAASFTIPANSDAFLPCEAVGNPLPTIHWTRVS-GL
mouse_5_corrected	-----
rat	EISQGTQKSRFHVLPNGTLSIQRVSIQDRGQYLCSAFNPLGVDHFVSLSVVFPARILD
human_5+3_corrected	DLSRGNQNSRVQVLPNGTLSIQRVSIQDRGQYLCASNLFGTDHLHVLSVVSYPFRILE
mouse_5_corrected	-----
rat	RHVKEITVHFGSTVELKCRVEGMPRTVSWILANQTVVSETAKGSRKVVWTPDGTLIYN
human_5+3_corrected	RRTKEITVHSGSTVELKCRABGRPSPTVTWILANQTVVSESSQGSRAVVTVDGTLVLHN
mouse_5_corrected	-----
rat	LSLYDRGFYKCVASNPSSGQDSLLVKIQVITAPPVITIEQKROAIVGVLGGSLLKLPCTAKGT
human_5+3_corrected	LSIYDRGFYKCVASNPSSGQDSLLVKIQVIAAPPVILEQRRQVIVGTWGESLLKLPCTAKGT
mouse_5_corrected	-----
rat	PQPSVHWVLYDGTTELKPLQLTHSRFTLYPMGTLYIRSIAPSVRGTYECIATSSSGSERRV
human_5+3_corrected	PQPSVYVWVLSDGTVEVKFLQFTNSKLFVFNGLTYIRNLASSDRGTYECIATSSSGSERRV
mouse_5_corrected	-----
rat	VILTVEEGETIPRIETASQKTEVNLGEKLLNCSATGDPKPRIIWRLPSKAVIDQWHRM
human_5+3_corrected	VMLTMEERVTSPRIEASQKTEVNFEGDKLLNCSATGEPKQIMWRLPSKAVVDQ----
mouse_5_corrected	-----
rat	GSRIHVYPNGSLVVGSVTEKDAGDYLCVARNKMGDDLVLHVRILRLTPAKIEQKQYFKKQ
human_5+3_corrected	GSWIHVYPNGSLFIGSVTEKDSGVYLCVARNKMGDDLILMHVSLRLKPAKIDHKQYFRKQ
mouse_5_corrected	-----
rat	VLHGKDFQVDCRASGSPVPEVSWSLPDGTVLNNVAQADDSGYRTKRYTLFHNGLTYFNNV
human_5+3_corrected	VLHGKDFQVDCRASGSPVPEISWSLPGDTMINNAMQADDSGHRTTRYTLFHNGLTYFNNV
mouse_5_corrected	-----
rat	GNAEEGDYICSAQNTLGKDEMKVHLTVLTAIPRIQSYKTTMRLRAGETAVIDCEVTGEP
human_5+3_corrected	GVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIQSNKTNKRIRKAGDTAVLDCEVTGDP
mouse_5_corrected	-----
rat	KPNVFWLLPSNNVISFSNDRFTFHANRTLSIHKVKPLDSGDYVCVAQNPSSGDDTKYKLD
human_5+3_corrected	KPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSGEYVCVARNPSSGDDTKMYKLD
mouse_5_corrected	-----
rat	IVSKPPLINGLYANKTVIKATAIRHSKYPDCRADGIPSSQVTWIMPGNIFLPAPYFGSR
human_5+3_corrected	VVSKPPLINGLYTNRTVIKATAVRHSKHFDCRAEGTPSPEVMWIMPDNIFLTAFYFGSR
mouse_5_corrected	-----
rat	VTVHPNGTLEMRNIRLSDSADFTCVVRSEGGESVLVQLEVLEMLRRPTFRNPFNEKVIA
human_5+3_corrected	ITVHKNGTLEIRNVRLSDSADFCVARNEGGEVLVQLEVLEMLRRPTFRNPFNEKIVA
mouse_5_corrected	-----
rat	QAGKPVALNCSVDGNPPPEITWILPDGTQFANRPNSPYIMAGNGSLILYKATRNKSGKY
human_5+3_corrected	QLGKSTALNCSVDGNPPPEIIWILFNGTRFSNGPQSYQYLIASNGSFIISKTTREDAGKY
mouse_5_corrected	-----
rat	RCAARNKVGYTEKLILLEIGQKPVILTYEPGMVKSVSGEPLSLHCVS DGIPKPNVKWTFP
human_5+3_corrected	RCAARNKVGYIEKLVILEIGQKPVILTYAPGTIVKGISGSLSLHCVS DGIPKPNIKWTFP
mouse_5_corrected	-----

rat GGHVIDRPQVDGKYILHENGTLVIKATTAHDQGNVICRAQNSVGQAVISVSVMVVAIRK
human_5+3_corrected SGYVVDRPQINGKYILHDNGTLVIKEATAYDRGNICKAONSVGHTLITVPVMIVAYPPR
mouse_5_corrected -----

rat IINYLPRNMLRRTGEANOLHCVALGIPKPKVTWETPRHSLLSKATARKPHRSEMLHPQGT
human_5+3_corrected ITNRPPRSIVTRTGAAFPQLHCVALGVPKPEITWEMPDHSLSTASKERTHGSEQLHLQGT
mouse_5_corrected -----

rat LVIQNLQTSDSGVYKCAQNLGTDYATTYIQV
human_5+3_corrected LVIQNPQTSDSGIYKCTAKNPLGSDYATTYIQV
mouse_5_corrected -----

(rat: SEQ ID NO:10)

(human_5+3_corrected: SEQ ID NO:11)

(mouse_5_corrected: SEQ ID NO:12)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Figure 16

```

rat      MOVVRGREGVSGLLISLTAVCLVVTTPGSRACPRACACVVFTEVHCTFRYLTSIPDGIIPANVE
human_5+3_corrected MKVKGRGITCLLVSFVAVICLVATPGGKACPRRCACVMPTEVHCTFRYLTSIPDSIPPNVE
*:*** :: **:***:****:*****:*****:*****:*****:***

rat      RINLGYNSLTRLTENEDFDGLSKLELIMLHNSNGIHRVSDKTFSGLQSLQVLKMSYNKVQII
human_5+3_corrected RINLGYNSLVRIMETDFSGLTKELELIMLHNSNGIHTIPDKTFSDLOALQVLKMSYNKVRKL
*****:*** ***:***:*****:****:*****:*****:*****:

rat      RKDTFYGLGSLVRLHLDHNNIEFINPEAFYGLTSRLVHLEGNRLTKLHPDTFVSLSYIQ
human_5+3_corrected QKDTFYGLRSLTRLEMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYIQ
***** ***:***:*****:****:*****:*****:*****:

rat      IFKTSFIKYLFLSDNFLTSLPKENVSYMPNLESILYHGNPWTCCHLKWLSWMQGNPDI
human_5+3_corrected IFKISFIKYLFLSDNFLTSLPOENVSYMPDLDSLYHGNPWTCCHLKWLSWMIQ--DIV
** ***:***:*****:*****:*****:*****:*****:***

rat      IKCKKDRSSSSPQQCFPCNPRISKGRFFAMVPSGAFLCTKPTIDPSLKSLSLTQEDNG
human_5+3_corrected IKCKKDRSPSSAQCFPCNPRTSKGRPLAMVSAAPQCAKPTIDSSLSKSLSLTILEDSS
*****:***:*****:****:***:****:*****:****

rat      SASTSPQDFIEPFGSLSNMTCKSGNKMVCSIQKPSRTSPTAEENDYIMNASFST
human_5+3_corrected SAFISPOGFMAPPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPLAEENDYIVLNTSFT
** ***:***:*****:****:***:*****:*****:*****:***

rat      NLVCSVDYNHIOQVWQLLALYSDSLILERKPOLTFETPSLSRYKQVALRPEDIFTSIEA
human_5+3_corrected FLVCNIDYGHIOQVWQILALYSDSLILERSHLLSETPOLYKYKQVAPKPEDIFTNIEA
***:***:*****:*****:****:***:*****:***

rat      DVRADPFWFQQEKIVLQLNRTATTLSTLQIQFSTDAQIALPRAEMRAERLKWIMLMNN
human_5+3_corrected DLRADPSWLMQDQISLQLNRTATTFSTLQIQYSSDAQITLPRAEPRVVKHWTMISRDN
*:*** *:***:*****:*****:***:*****:*****:

rat      PKLERTVLVGGTIALSCPGKGDPSPHLEWLLADGSKVRAPYVSEDGRILIDKNGKLELOM
human_5+3_corrected TKLEHTVLVGGTVGLNCPGQDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELOM
***:*****:*,***:***:***:*****:*****:*****:

rat      ADSFDAGLYHCISTNDADADVLTYRITVVEPYGESTHDSGVQHTVVTGETLDELPLSTGV
human_5+3_corrected ADSFDGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIHHTVFTGETLDELPLCHSTGI
*****:***:***:*****:***:***:*****:*****:

rat      PDASISWILPGNTVFSQPSRDEQILNNGTLRLQVTPKQGHYQCVANPSCADFSFVKV
human_5+3_corrected PDASISWVPGNNVLVQSSRDEKVLNNGTLRLQVTPKQGHYRCVANPSCGVDFLIFOV
*****:***:***:*****:*****:*****:*****:

rat      SVQKKGQRMVEHDREAGSGGLGEPNSSVSLKOPASLKLASALTGSEAGQVSGVHRKNK
human_5+3_corrected SVKMKGQRFLEHDETEGSGLDESNIPLHLKEFPFGAQLRTSALMEAEVGGHTSSTSKRHN
**:* ***:*** *:***:***.***.***:***:***:***:***:***:

rat      HRDLIHRRGDSTLRRFREHRRQLPLSARRIDPQWAALEKAKKNSVPKKQENTITVKEV
human_5+3_corrected YRELTQRRGDSTHRRFREHRRHFPSSARRIDPQWAALEKAKKNAMPKRENTITVSEF
*:** :*****:*****:***:*****:*****:*****:***:*****:

rat      PLAVPLVELTDEEKDASGMIPPEEFMVLTKASGVFGRSPTADSGFVNHGFMSTIASGT
human_5+3_corrected PVVTQLPNIPGEEDSSGMLALHEEFNVFATKALNLPARTVITADSTISDPMNTINXGT
*... * :...*:***:***:***** ***:***:***:***:***:***:

rat      EVS-TVNPQTQGEHLPDFKLSVTNGTAVTKSMNPSIASKIEDTINQNPILIF---SV
human_5+3_corrected EFSFVNSQILPPEEPTDFKLSLTAIKTAMSKNINPTMSSQIQGTINQHSSTVFLLGA
**.* ***:*** ***:***:***:***:***:***:***:***:

rat      AEIRDSAQAGRAS--SQSAHFVTGGNMATYGHNTNTYSSFTSKASTVLQPINPTESYGPQI
human_5+3_corrected TEFQDSQMGREGHEFQSRPPITVRIMIKDVNVKMLSSTNKL---LLESVNTTNSH--QT
*:*** ***:*** ***:***:***:***:***:***:***:***:

```


(rat: SEQ ID NO:13)
(human 5+3 corrected: SEQ ID NO:14)

Figure 17

MQKRGREVSCLLISLTAICLVVTPGSRVCPRRCACYPTEVHCTFRDLTSIPDGPANVER
 VNLGYNSLTRLTENDFSGLSRLELLMLHSNGIHRVSDKTFSGLQSLQVLKMSYNKVQIIE
 KDTLYGLRSLTRLHLDHNNIEFINPEAFYGLTLLRLVHLEGNRLTKLHPDTFVSLSYLQIF
 KTSFIKXLYLYDNFTSLPKEMVSSMPNLESYLHGPNWTCDCHLKWLSEWMQGNP
 (SEQ ID NO: 15)

Figure 18

MKVKGRGITCLLVSFAVICLVATPGGKACPRRCACYPTEVHCTFRYLTSIPDSIPPVNE
 RINLGYNLVRMLMETDFSGLTKELELLMLHSNGIHTIPDKTFSDLQALQVLKMSYNKVRK
 LQKDTFYGLRSLTRLHMDHNNIEFINPEVFYGLNFLRLVHLEGNQLTKLHPDTFVSLSYL
 QIFKISFIKFLYLSDNFLTSLPQEMSYPDLDSLLYLHGPNWTCDCHLRWLSDWIQPDVI
 KCKKDRSPSSAQQCPLCMNPRTSKGKPLAMVSAAAFQCAKPTIDSSLKSKSLTILEDSSS
 AFISPQGFMAPFGSLTLNMTDQSGNEANMVCSIQKPSRTSPIAFTEENDYIVLNTSFSTFL
 VCNIDYGHQIPVWQILALYSDSPLILERSHLLSETPQLYYKYKQVAPKPEDIFTNIEADLR
 ADPSWLMQDQISLQLNRTATTFTSLQIQYSSDAQITLPRAEVRPVKHKWTMISRDNNTK
 LEHTVLVGGTVGLNCPGQGDPTPHVDWLLADGSKVRAPYVSEDGRILIDKSGKLELQM
 ADSFDTGVYHCISSNYDDADILTYRITVVEPLVEAYQENGIIHTVFIGETLDLPCHSTGIP
 DASISWVPGNNVLYQSSRDKKVLNNGTLRLQVTPKDQGYRCVAANPSGVDFLIFQV
 SVKMKGQRPLEHGETEGSGLDESNPPIAHLKEPPGAQLRTSALMEAEVKGHTSSTSKRH
 NYRELTLQRRGDSTHRRFRENRRHFPPSARRIDPQHWAALEKAKKNAMPDKRENTTV
 SPPPVVTQLPNIPGEEDDSSGMLALHEEFMVPATKALNLPARTVTADSRTISDSPMTNIN
 YGTEFSPVVNSQILPPEPTDFKLSTAILTTAMSKNINPTMSSQIQGTTNQHSSTVFPLLLG
 ATEFQSDSQMGRGREHFQSRPPITVRTMIKDVNVKMLSSTTNKLLLESVNSHQTSVREV
 SEPRHNHFYSHTTQILSTSTFPSDPHTAAHSQFPIPRNSTVNIPLFRRFGRQRKIGGRGRIIS
 PYRTPVLRHRYSIFRSTTRGSSEKSTTAFSATVLNVTCLSCLPRELTTATAALSFPSAAPI
 TFPKADIARVPSEESTTLVQNPLLLLENKPSVEKTTPTIKYFRTEISQVTPTGAVMTYAPT
 SIPMEKTHKVNASYPRVSSTNEAKRDSVITSSLGAIKPPMTIIAITRFSRRKIPWQQNFV
 NNHNPKGRLRNQHKVSLQKSTAVMLPKTSPALPQRQSSPFHFRRLSTSVMQIPSNLTLT
 AHHTTTKTHNPGSLPTKKELPFPLNPMLPSIISKDSSTKSIISTQTAIPATTPTFPASVITYE
 TQTERSRAQTIQREQEPQKNRTDPNISPQSSGFTTPTAMTPPALAFTHSPPENTTGISST
 ISFHSRTLNLTDVIEELAQASTQTLKSTIASSETLSSKSHQSTTTTRKASLDTPIPPFLSSSAT
 LMPVPISPPFTQRAVTDTRGDSHFRLMTNTVVKLHESSRHNLQMPSSQLEPLTSSTSNLL
 HSTPMPALTTVKSQNSKLTPSPWAEQFWHKPYSDIAEKGGKPEVSMLATTGLSEATTLV
 SDWDGQKNTKKSDFDKPKVQEATTSKLLPFDSLRYIFEKPRIVGGKAASFTIPANSDAF
 LPCEAVGNPLPTIHWTRVSGLDLSRGNQNSRVQVLPNGTLSIQRVEIQDRFQYLCASNL
 FGTDHLHVTLVSVSYPPRILERRTKEITVHSGSTVELKCAEGRPSPTVTWILANQTVVSE
 SSQGSRQAVVTVDGTLVLHNSIYDRGFYKCVASNPGGQDSLLVKIQVIAAPPVILEQRR
 QVIVGTWGESLKLPCAKGTPQPSVYWVLSDGTEVKPLQFTNSKLLFLFSNGTLYIRNLAS

SDRFTYECIATSSTGSERRVVMLTMEERVTSPIREAASQKRTEVNFQDKLLLNCSATGEP
 KPQIMRLPSKAVVDQGSWIHYPNGSLFIGVTEKDSGVYLCVARNKMGDDLILMHVSLR
 LKPAKIDHKQYFRKQVLHGKDFQVDCKASGSPVPEISWSLPDGTMINNAMQADDSGHR
 TRRYRLFNNGTLYFNKVGVAEEGDYTCYAQNTLGKDEMKVHLTVITAAPRIRQSNKTN
 KRIKAGDTAAVLDCEVTGDPKPKIFWLLPSNDMISFSIDRYTFHANGSLTINKVKLLDSG
 EYVCVARNPSGDDTKMYKLDVVS KPPLINGLYTNRTVIKATAVRHKKHFDCRAEGTP
 SPEVMWIMPDNIFLTAPYYGSRITVHKNGTLEIRNVRLSADFICVARNEGGEVLVVQLE
 VLEMLRRPTFRNPRNPFNEKIVAQLGKSTALNCVSVDGNPPPEIHWILPNGTRFSNGPQSY
 QYLIASNGSFIISKTTREDAGKYRCAARNKVGYYIEKL VILEIGQKPVILTYAPGTVKGISGE
 SLSLHCVSDGIPKPNIKWTMPSGYVVD RPQINGKYILHDNGTLVIKEATA YDRGNYICKA
 QNSVGHTLITVPVMIVAYPPRITNRPPRSIVTRTGAA FQLHCV ALGVPKPEITWEMPDHS
 LLSTASKERTHGSEQLHLQGT LVIQNPQTSDSGIYKCTAKNPLGSDYAATYIQVI (SEQ ID
 NO: 16)

Figure 19

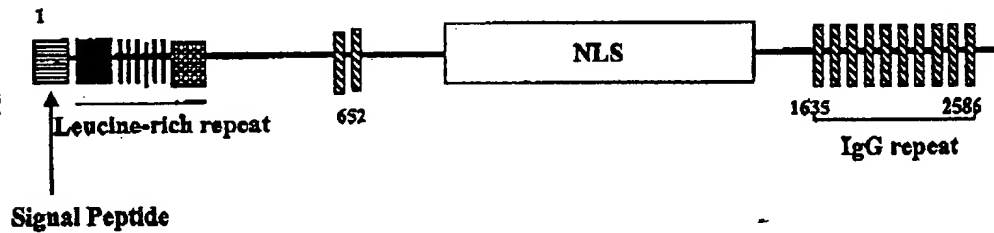


Figure 20

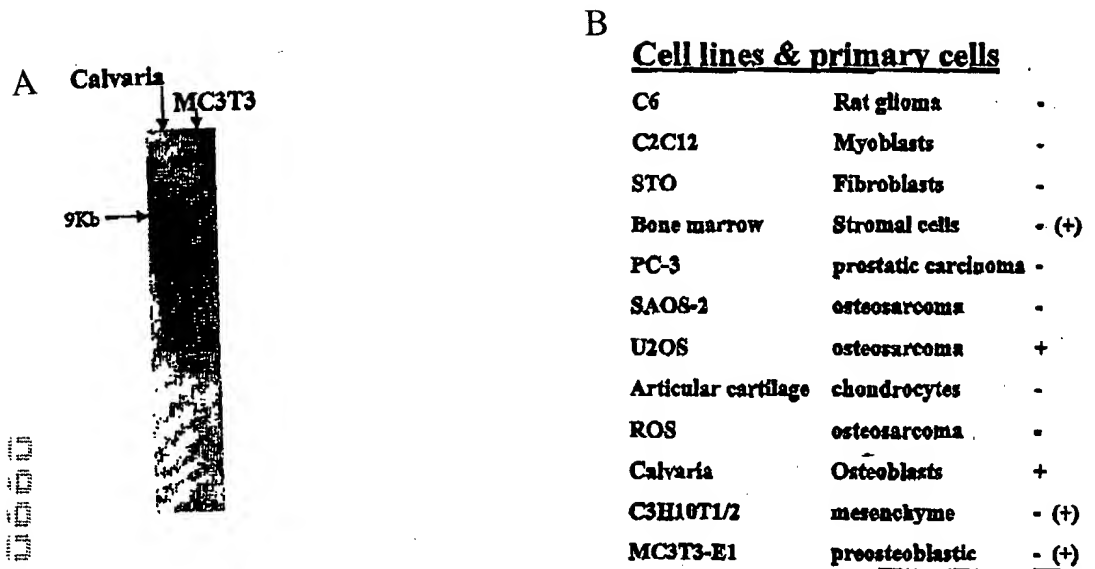


Figure 21

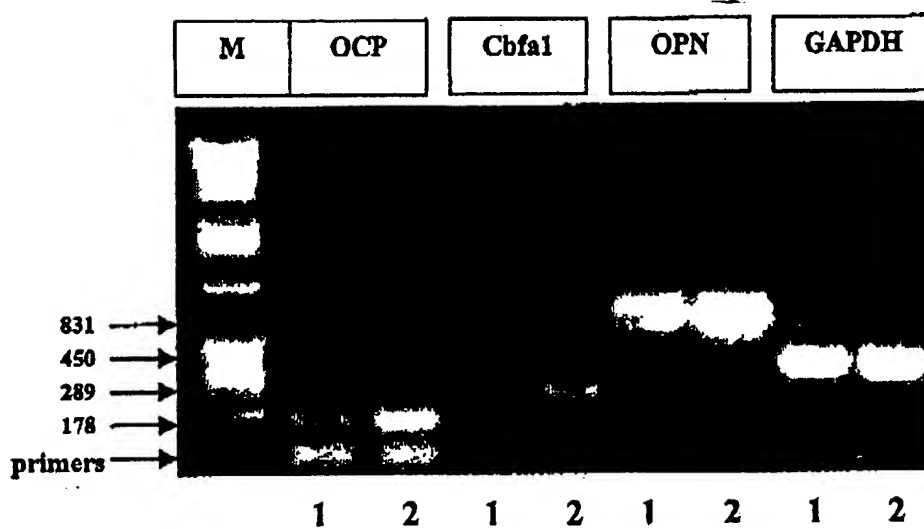


Figure 24

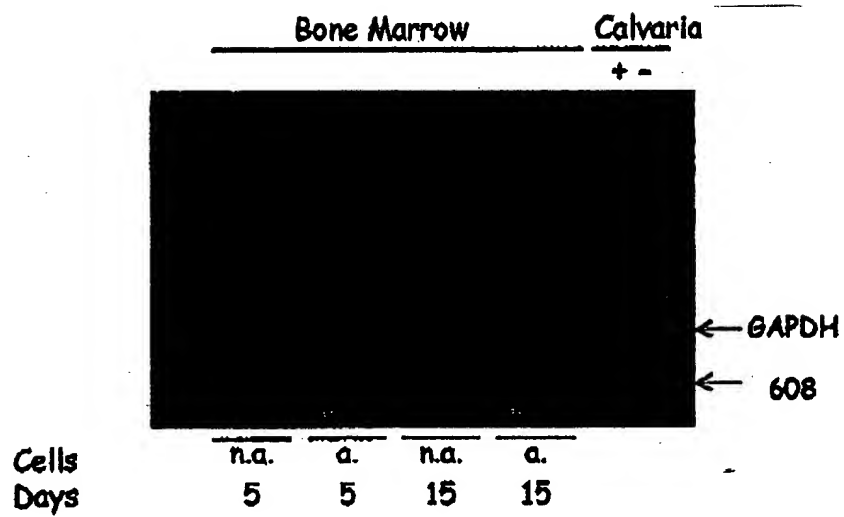
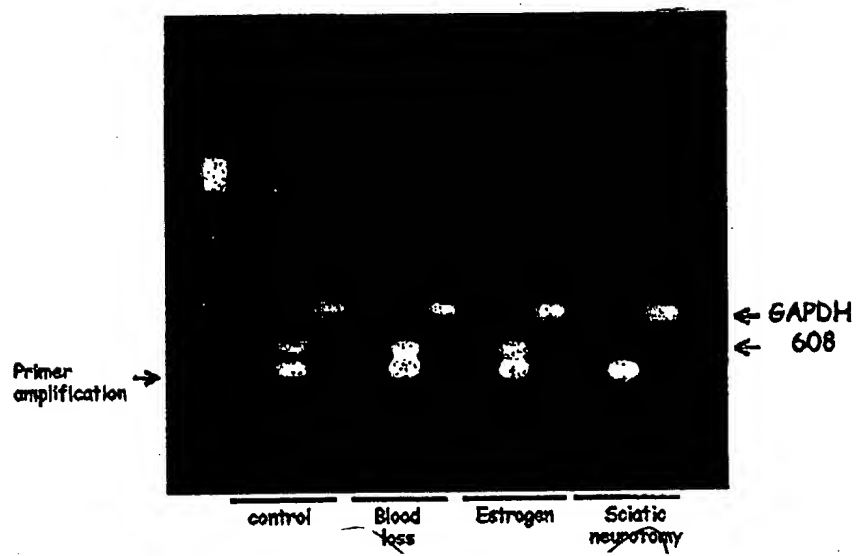


Figure 25



	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167	2168	2169	2170	2171	2172	2173	2174	2175	2176	2177	2178	2179	2180	2181	2182	2183	2184	2185	2186	2187	2188	2189	2190	2191	2192	2193	2194	2195	2196	2197	2198	2199	2200	2201	2202	2203	2204	2205	2206	2207	2208	2209	2210	2211	2212	2213	2214	2215	2216	2217	2218	2219	2220	2221	2222	2223	2224	2225	2226	2227	2228	2229	2230	2231	2232	2233	2234	2235	2236	2237	2238	2239	2240	2241	2242	2243	2244	2245	2246	2247	2248	2249	2250	2251	2252	2253	2254	2255	2256	2257	2258	2259	2260	2261	2262	2263	2264	2265	2266	2267	2268	2269	2270	2271	2272	2273	2274	2275	2276	2277	2278	2279	2280	2281	2282	2283	2284	2285	2286	2287	2288	2289	2290	2291	2292	2293	2294	2295	2296	2297	2298	2299	2300	2301	2302	2303	2304	2305	2306	2307	2308	2309	2310	2311	2312	2313	2314	2315	2316	2317	2318	2319	2320	2321	2322	2323	2324	2325	2326	2327	2328	2329	2330	2331	2332	2333	2334	2335	2336	2337	2338	2339	2340	2341	2342	2343	2344	2345	2346	2347	2348	2349	2350	2351	2352	2353	2354	2355	2356	2357	2358	2359	2360	2361	2362	2363	2364	2365	2366	2367	2368	2369	2370	2371	2372	2373	2374	2375	2376	2377	2378	2379	2380	2381	2382	2383	2384	2385	2386	2387	2388	2389	2390	2391	2392	2393	2394	2395	2396	2397	2398	2399	2400	2401	2402	2403	2404	2405	2406	2407	2408	2409	2410	2411	2412	2
--	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	---

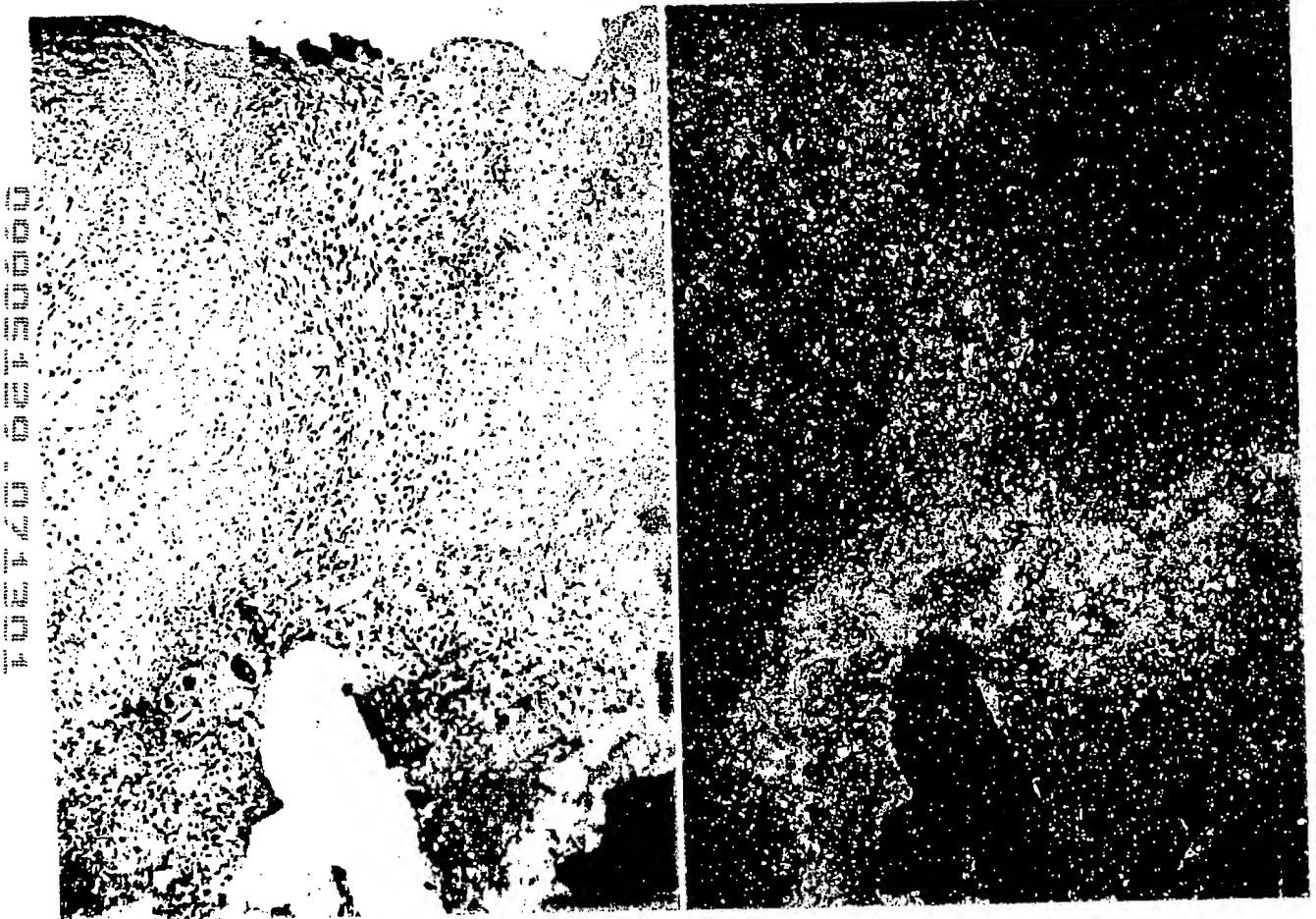


54/85

Figure 27

A

B

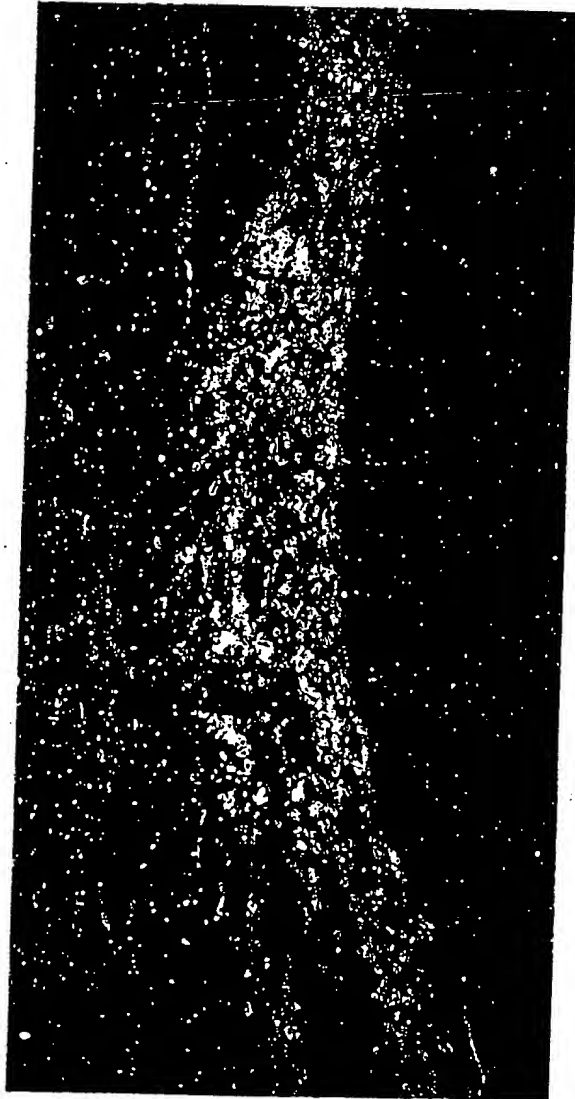


Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	



1

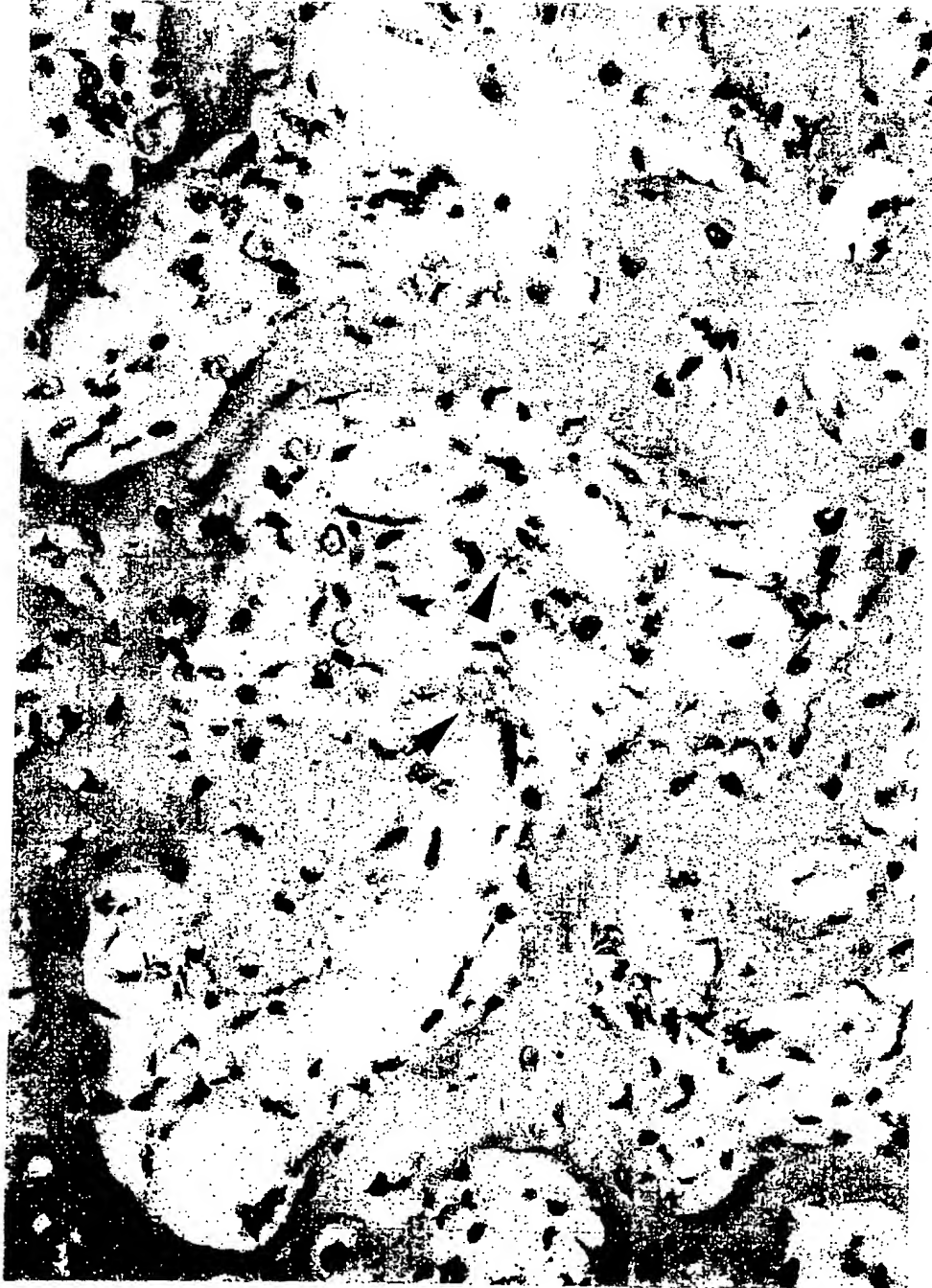
B



Year	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

57/85

Figure 30



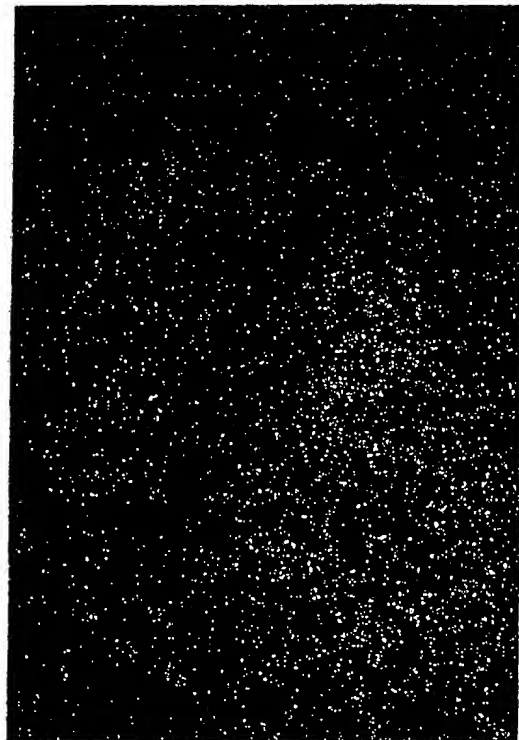
00005430 074304

[illegible]

A



B





This image shows a highly textured, high-contrast surface, likely a book cover or endpaper. The background is a light, grainy white, densely populated with small, dark, irregular spots and fibers. These dark elements are scattered across the entire frame, creating a mottled and speckled appearance. The overall effect is one of extreme contrast and a rough, tactile quality.

61/85

Figure 34

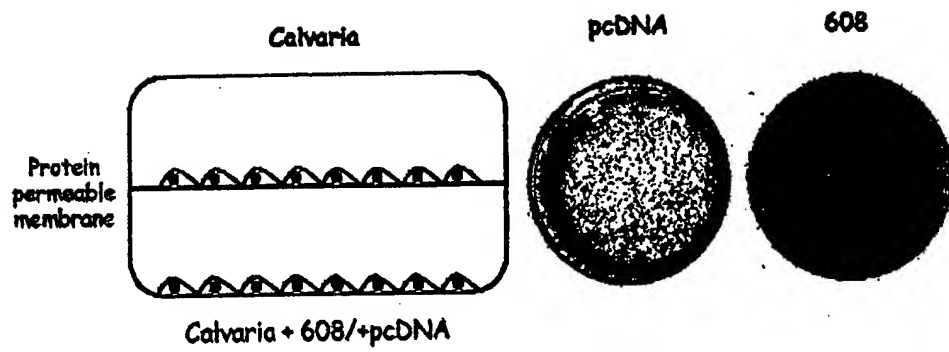
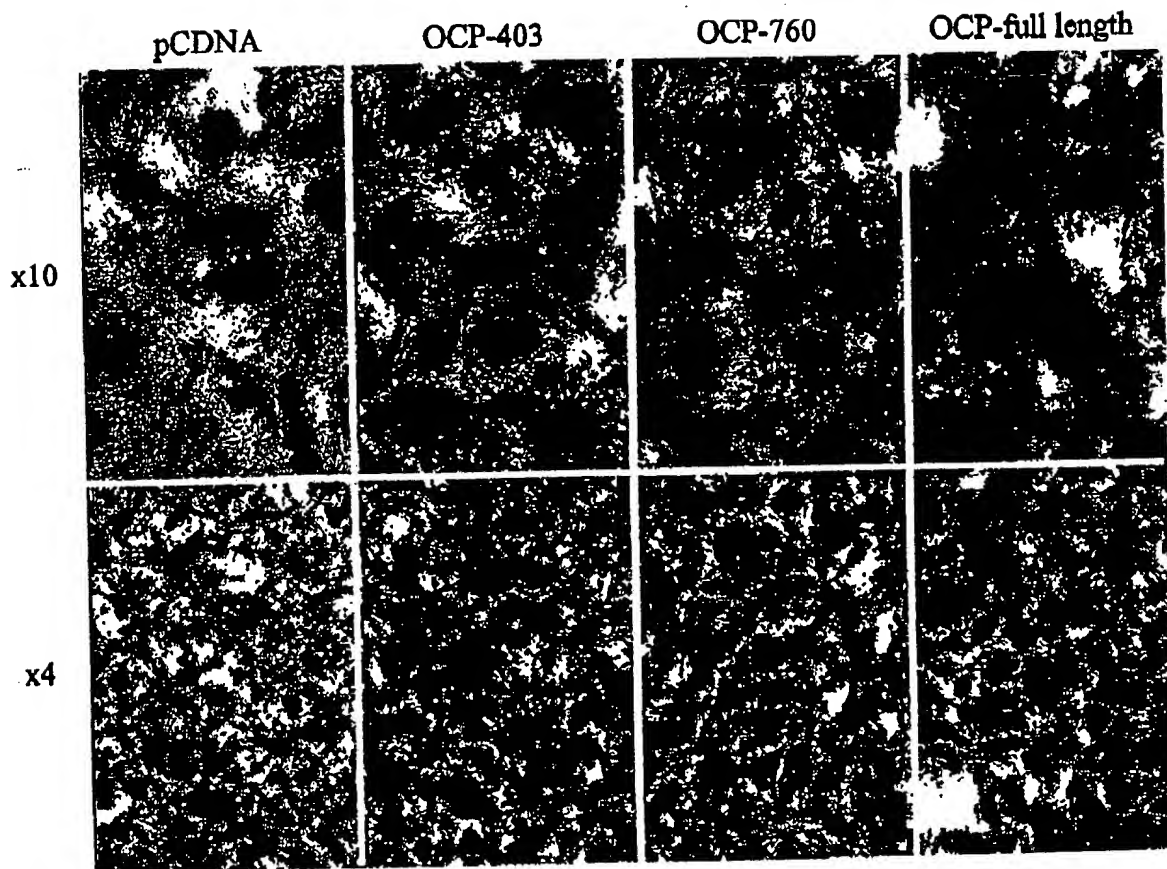


Figure 35



62/85

Figure 36

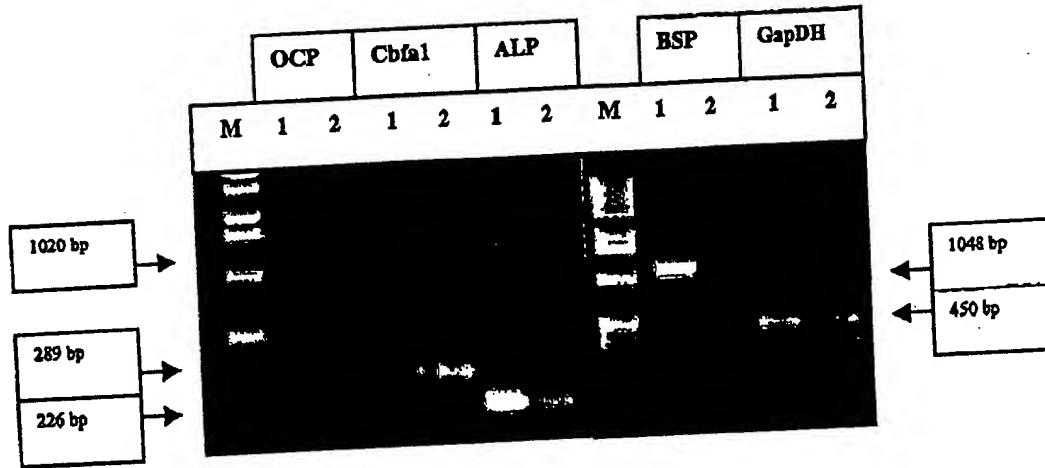


Figure 37

x4

x10

pCDNA
ROS stable line

OCP
ROS stable line



Figure 38

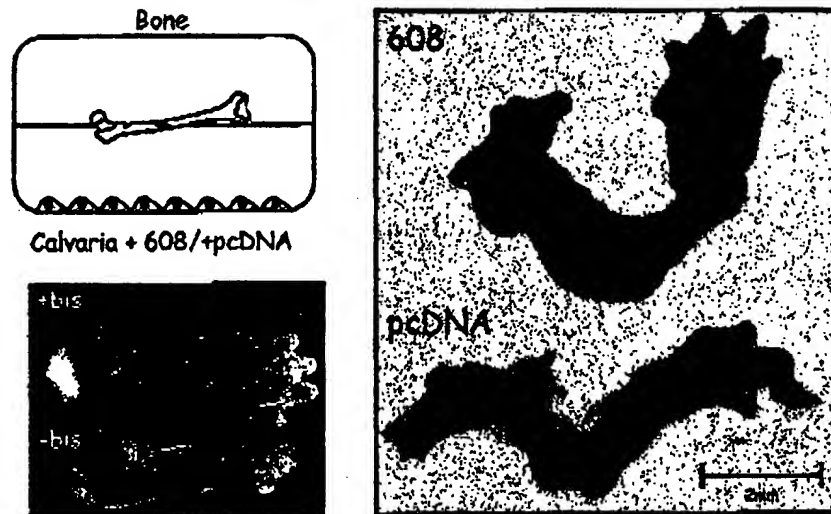
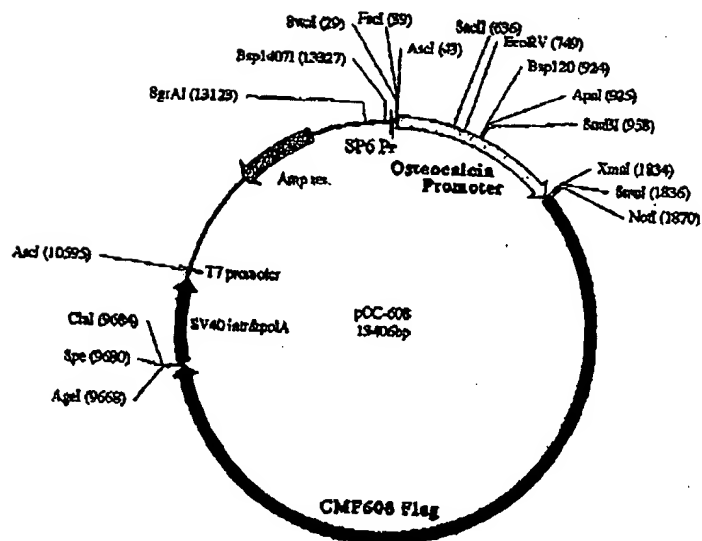
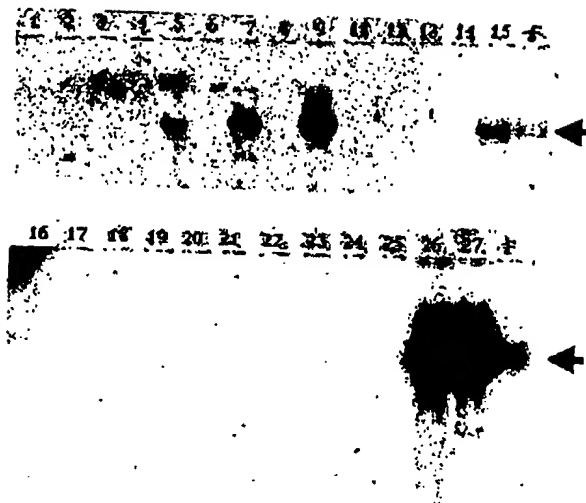
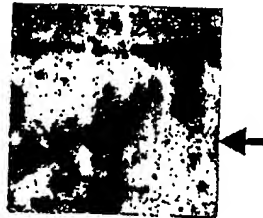


Figure 39



A

**B**

C

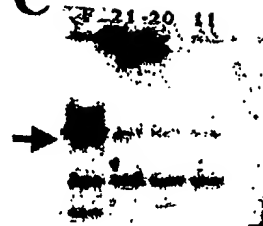


Figure 41

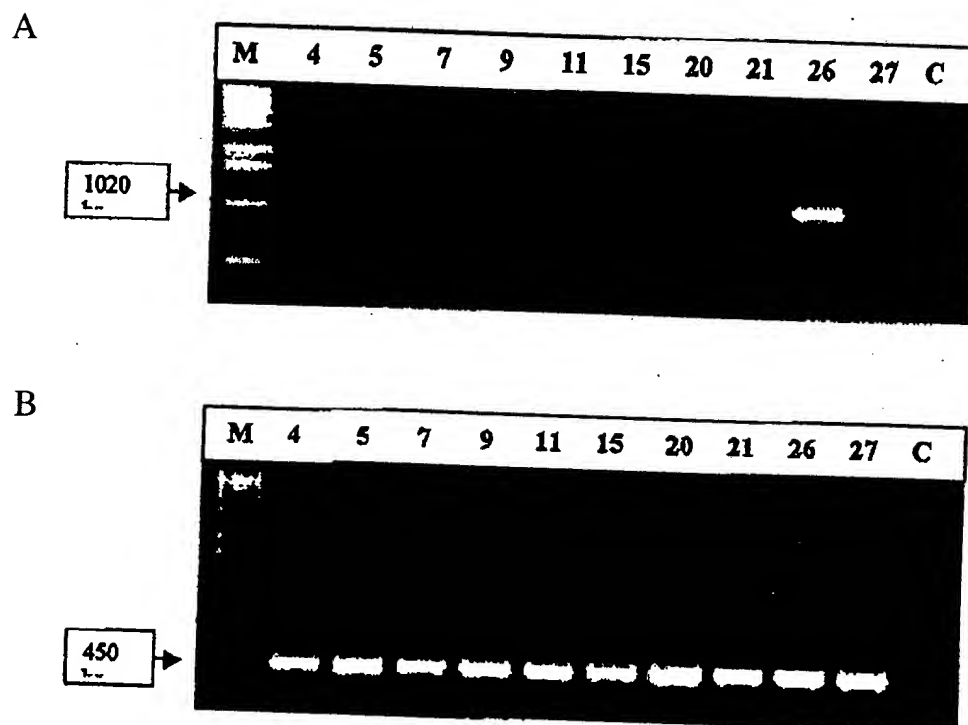
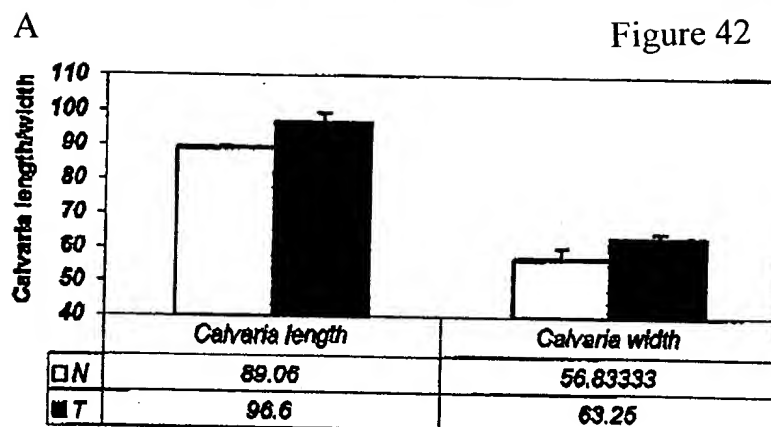


Figure 42



B

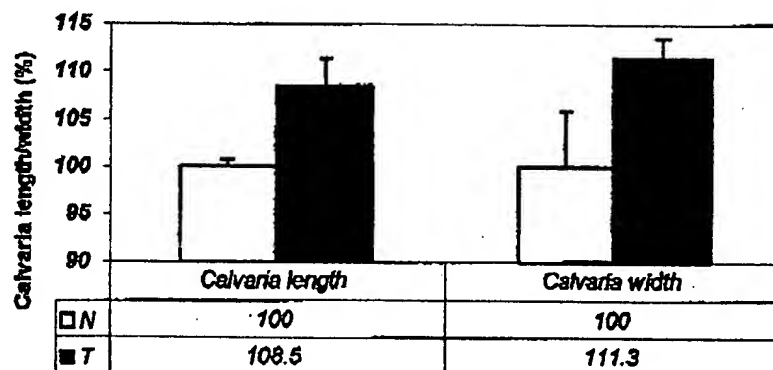


Figure 44

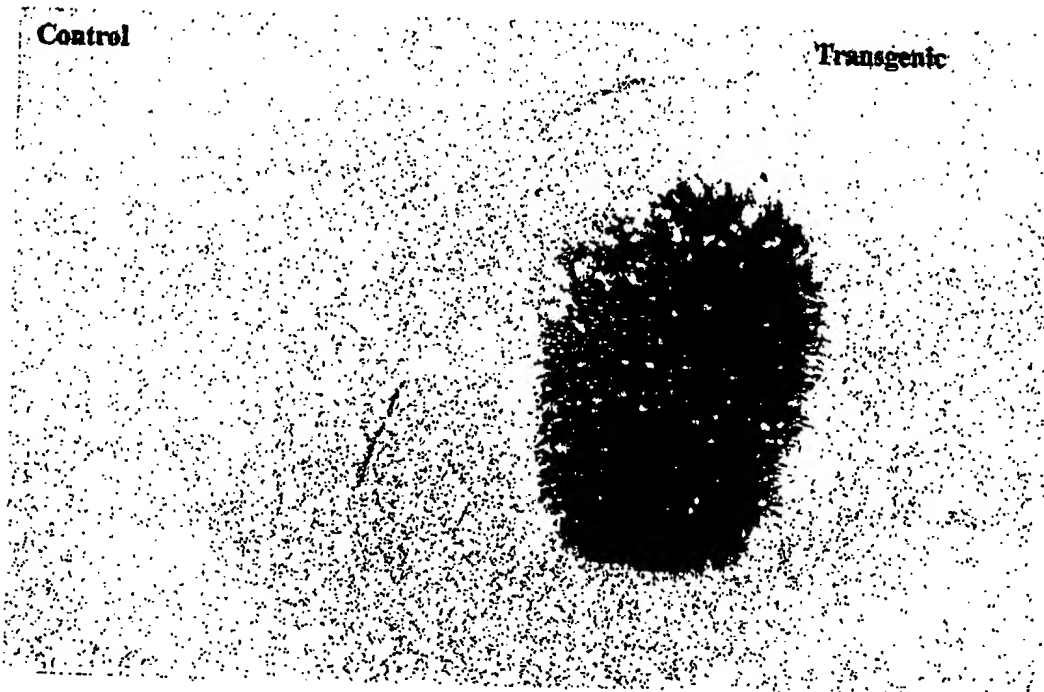
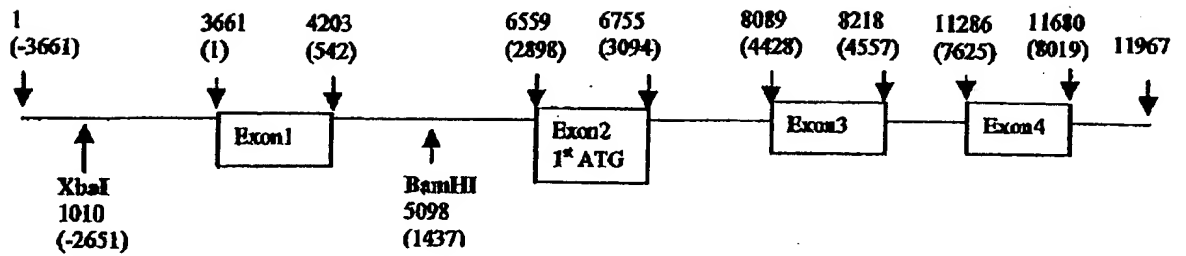


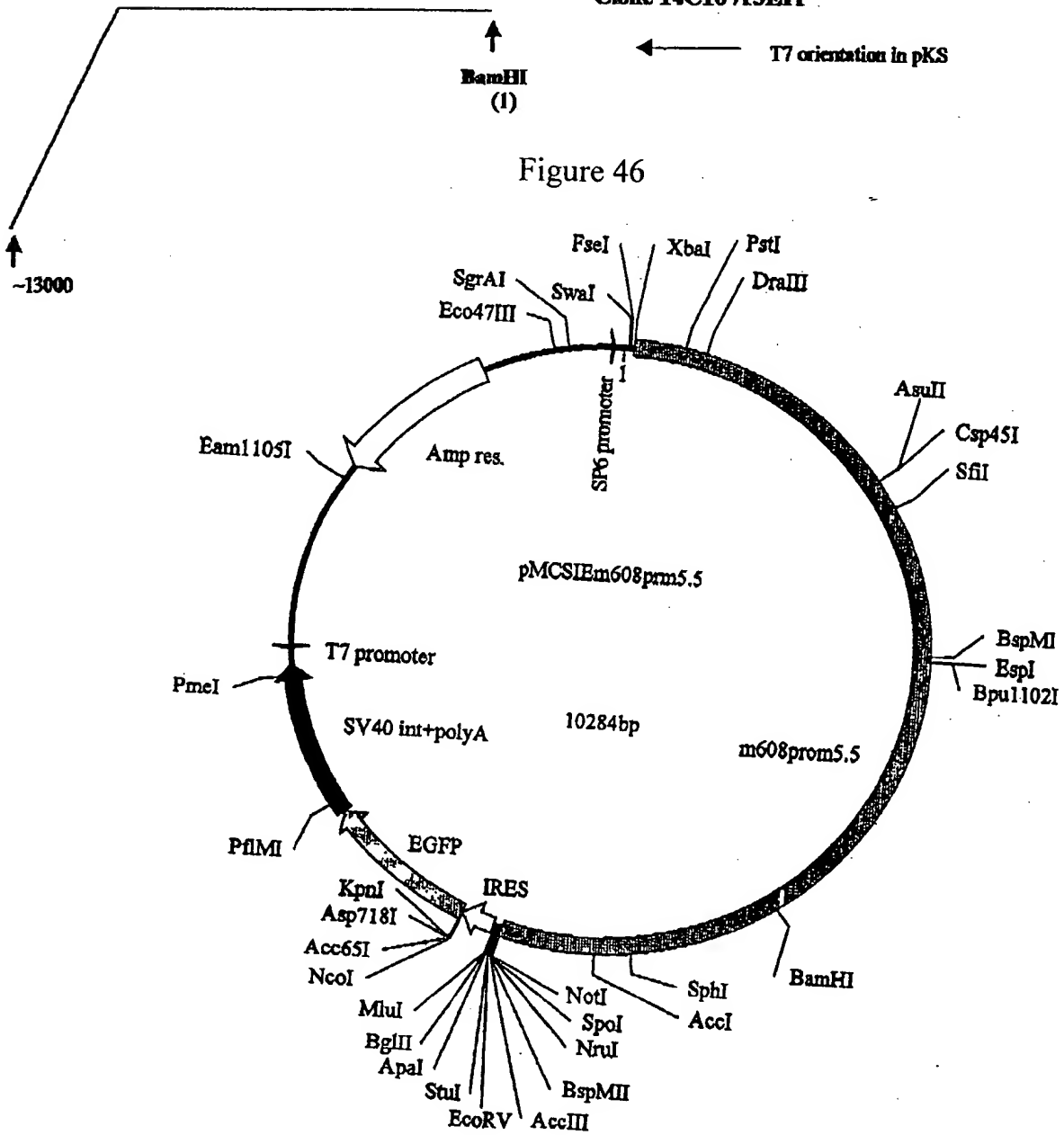
Figure 45



Clone 14C10 /15E11

T7 orientation in pKS

Figure 46



10284bp

Figure 47

↓ (XbaI)

TCTAGAAGTAAATGATCCTGAGTAGCGATCCTGGGAAAATACGTA
CTAACACACTGC
AATCATCTCTCTGTGGTTTGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATA
ACCTACCGAATACAGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCT
GTAGCTGCCAGAACATACTAGATGCCATATTTCCAAGGCAAGTGCCACATGCGGACAT
CTTAAGAATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTTCGTGCCACACAGGTCTCCAG
GGTTAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTACCTTGATCATTAAATGACTAGAA
ATCACTTGGTGATTAGCAACTGGATATGGAATATTACTAATTTGTACCCAAAGCCAGGCCAC
CTCAGCTTTGGCAGCTCCATTCTCTGTGGAGCCCAGTCACGTGGGTTTGAATCAACTGT
ACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTCTTCAAGCTCTTTT
GAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAGTGAGAGCAGCATGAATG
TGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAATAAACTGACTTTAAAAA
AGCTTATGATACTTGTACAGAGTAAATCTTCCATAAATATCATCTGCATTTATAAATTAT
TTTCATAATCCATCAATTAAAAACCTTTAGAAATTTTGTTAACACAAAGATCCCTAGGCC
CTGCCCTAGGATGGTCTGTATGGTGGGCCCTGAGAGATGGAGCTTAAGA
ACTTACTTGCTCC
AGGAGCACATCTTCAGAACATCTGCCTCAAAACATTTATCCCAAATGCTCATCAAAAGGCTC
ACTCACATGTGCTTCAACACAGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGT
GGAAAGCCTGCTAGAGGAACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTC
ATAACCCCTCACCACAACAAACAAACGACAACAAACCCATAAAAAATTATCACGGCAAAAT
GAACTAAGCCATATGCAGAAAAAGTATTATATGTTCTCATTTGTGGGGTGTTTTCTTAAAT
AGTCAAAATATGCAGAAATATAGACAAAGATGGTTATGCAAGTGGGGATGGCGAAGGATA
CTTGATAGATTAGAGGACACAAAGCAACAACTACAGAGTGAAGTAATCCAGAGACTTAATG
TATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAAACGAGTGTATCTT
ACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGTATCTTCATG
TAGCTGATAATTTATATTGTACACCTCAACATAGATAACCAACAAAGAGGAAGAGGAT
AGGTGCCCTCTCCAGGGCGGAAGAGTACATTCGAAAGTCAGACACCATTTGTGTAGATGTA
CCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAGGGATCTGCAACCCATAGG
TGGAACAACATTATGAGCTAACCAAGTACCCCGGAGCTCTTGACTCTAGCTGCATATATATC
AAAAGATGGCCTAATCGGCCATCACTGGAAAGAGAGGCCCATTTGGACTTGCAAACTTTAT
ATGCCCCAGTACAGGGGAATACCAAGGGCCAAAAAGGGGGAGTGGGTGGGCAGGGGAGTG
GGGGTGGGTGGATATGGGGGACTTTTGGTATAGCATTGGAAATGTAAATGAGTTAAATAC
CTAATAAAAAATGGAAAAAAGGAAAGGAAAGGTCAGACACCTCAC
TTCACTGCTATCTCAACTTGCAAAACAGAAGGGGAGTCACAAACCCAGGACAAACACAGT
GATTGAAGCGTCTTTGAATGTTATTGCTGTGTTGTTTACCACCATCATTAGCATATATTCAT
TGTGAAAACCTACGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTACGC
ATATTTGGCAACCTACAGCCCGAGCCCTTTGCCCCGCCCCCAACACACACACACACAC
CACCTTTACCTTCTCCTGGG
CATCATCTGCTCACTCAACCAACCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTCTC
CTATTTTATGTCTAGGTTCCCTTCCCTGTTAGGAGATGGGAGAGGTCACGAAAGGAA
AGAATTTGTAGCCCTGAGCCAGCCCGGCCACAGAGCCTGCCACAGACAGGAAAAAGCC
CAGGGCTTACCAGCACAGGAGGAGCAAACTCGCAGGCGAGCCTGGGTGGCGCTGGTGGT
CCCGGGTCGATGGCCCGCCATTTCCAGAAGCCGAGGCTATAGCTGCGTCACCTGCCCGG
CCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGACCCCAAGGCGGTGCAATTCCA
TTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCCTCGGGAAGTGAAGGGACCCAGAG

Exon1

AGTCTGGTAGATGTGGGAGCTGGGGTTACAGGGCGAGACAGAGGGTGGGATGGGCAGAAG
GGTCCAGGAAAAAGGAAAGTACTGGAGGGGAGTTGGGACAAAAGCAGCGACCAAGGGA
CATCGCTTCAGTGACTGAAGCCAGGCAAAAGGAGCGGGAAGGATTATGTAGCTGGGA
CGCTTTCATAAACTGATGACGTGTTTGTGCAAGCAAGCAATTTAGGAGAAACGCCT
GGGACGTGGGAAAGAAGGAGTGATCGATTAGTACTTGTAAAGTTTGGTGAAGTTGAGAAC
TAACCTATACTATTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAGC
GCAATCAGATAAAGGAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAATAACA
GAAGCGCCTCCAGTCTCTGAAGAGTCAGTCCCCAGCTAGTGAAGACTAAGCCTACTAA
GCCTTTTGTCTCCCTTGGAAAGCAAAAGCGTTCTTCAATCAGGTGAAGGCTCTCTCAGA

AGATTTCTCTCTCTGCTTATGTTACAAGAGGATTCAAAAAGCAAGACAGAAGAGCTCAGG
 TATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACATCATGTGGTA
 TTCATTTACGGCTCTCTTCTCTTATACTAACTCTTAAGGTGCATATAGTCTCTTCTGTTTCC
 CAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCTGCTTTTAAATCATC
 ACGCAGAGAGTATTCAAAAATATTCAGTGTATGTAACAGTGACAGTGTAGGCATAGAAAGTA
 ATCATTAGTAAATCTTAATTTGGGTTAACTCATTTCATAACAGCTCCAGGTTGGGAGGGAT
 CACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTTTCTAACAAGAGAAGCAGAATTCTTC
 CTGGCCATGCTCCCCATCACTGTGTCAAGTAAAGCAGAGGGGTGTTTCCAAGCAGAGAAAG
 AGCAGACAGTGTATGCCTGCAAAGTCAGAGACTCAGCCCTCCAGCTGGTCAGTTTACT
 GTCTCCCGGTCATTAGTTGGCTCTGAAAAGGCCCATGTGTCTTATTGGCAAGGACTTGC
 AGACATGCTAGAAAGAAATTTGACCTTTTTTCTAGTGGGTTATTACAGCTGTAAAAGTAT
 TTTGGAAGGTTAAGCCAAATAAATAAACACATATTAATAATACAATGTTTACAAAAT
 GATCATATAAAGAAGTACATTTCATAAATGCAATGTGAAAAATATATATAATTTTATCTAT
 TTAAGTGTGCAAAGTTTCTAAATTCACATGTACTATTTTATATTTATAAAAAATATTTT
 AAAATGTATATAAAGTGTAAGGCTCTTGGTCAAAACAAGAGAGTTAAATTTACAACT

↓ (BamHI)

TTAATTGTCCCGATAACATTATTATGATCTCTAATGACAGGATCTCTGCTTTTCATTGGGA
 AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCAAATG
 GGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCTGAGTCTATGCAAGTTTACA
 CTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTGTTGTGTCATCTTTATC
 TTCTTGATGATGAGCATTATAAGTGAATAAGATAGGATCTCAAAGGAATGTCAATTTGG
 ATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCACTAGTCTATTCAATTTATTGGATAA
 TTGGGGGATGGTGTTAATTTTTCAGTTCTTATGGAATTCCAAAAACAAAAACAAAC
 AAACAAACAAAAACCTCTGAACTAGAACTACCAATCCATTACTGGGTATGTAACAAAG
 AGAAATCTGCACAGAATTTATTGCTACATTGTTCAATTATTCACGACAGCCAAGAATGTGGA
 ACCAACTTACGTAGCCGTCAAAATATGAACGGATAAAGAAAAATGTGGAAATGTGTACAAC
 AGAGTCCCATGTGGCCATAAAAGAGTGAAATCATGACATATGCAGGAAATGGATGCAACT
 GGAATCAATTTGGGCTAATCAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTT
 CTCTGACAAACAGAAGCTAGATTTACACTTGTACGTGCGCATGTGTGTTAGAATTTTATT
 TAGTTATACACTATTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACA
 AGCTAGCTGGGGTGGGTAGGAGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
 GTAACATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGA
 GGCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAAAGT
 TTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAGGGGGCGTTCCAGTTT
 CTCCAGTCTGACACTGTGTATAACAACAGTTGACAATACAAAGTTGGCATGATGGATGG
 TTTTGTGCTATTTTTCATTTTCTTACTGTTTGTGTTGTGTTGTTGTTGTTGTTGTTGTTG
 GCTGTGTTTTCATTTGTTTCTTTGAGAGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATA
 AGCTGGAACGATCTGGAAGAAGTTGGGGAAAGAGAAAAATGTATGGAGCATATTTAA
 ACAAACAAACAAACAAACAAAGGTTCAATTTGCCACAAACAAAGGTGTGAATTAATTAAC
 CAGTTACGACTCTTAAAGAAATATTTCCCAATTATTTCCAGAGTTGCTATGTATGCTGTGC
 TAGGACTTTGCTTGAAGTGGCCCTATAACTCTGGTGTGGTGTCTTTTCAGG - Artificial
 NotI site

(SEQ ID NO:17)

CACAGACCTTCTCTCTCTAACTCTCTCCCCCATCTTGTGTGCTTCATCCCAGACTTCAACAC
CAGCAAGCACACTCTGCTAATGCAAGGGCTGCTCCTGTGCAGGACAACAAGGAGGCTGAAG
GCAGACCCACACGTTTCCAACTGCTCTCTGAGAGTCAATCCCCCTAGACTCATCTATAGCAG
GAAACCTGCTGTGATCTCCATTTCTTCTGTGACCACTCCCCAAGTTATCACAAGGAGTTT
TCTTCAAACCTTTCTCTCCAGCAAAACCCCTTCAGCTCCTTGGGTACTTTCTCTAGCCCTT
TCATTGGGAACCTGTGTCTCCATCCAAATGGATGGCTGTGAGCATCCACTTCTGTATAGAAT
CTTGGTCAGTGCAGTCTTTTGTATCCTCAAGAACTGGGTCTGAAAATTTTAACCCAAAG
AACTGTTTTTTGTTATGATTGCTGCAATCTCTTTCAATTCCAATAAAGAGTAAGCATCTCAT
TCCTTTGTCTCCTCCTTTTCAGTACCACCTGCGCTTTGTCTGCCTTTCTCAAAGAAATCAATAAA
ACCAAAGTGATATAGATTACATGGCATTCTCTAACTGCTACATCCACTCCAGTAGTATCTC
ACTTGGCAGGTGTAAAAGCCTGTGAAGCAGTCAAGGAGCATTTTCACAGAACTTAGCTCTC
CTGGAACCTTGGCATTCCCATAGCTAGAATGCCCGAGATTGTGCTGTAGATATTGTGGTG
GGTCTTGCATGCTTTCTTGCAGTATTTTACTGGATAAGAGTTAGAAATCTCAGGGCGAGCT
TAGCAAAAGTATACCTAGAATCTTCATGACAGTCAGGTATTGCAAACTACATTGCATATTA
GAAGAAAGTTGGTAAATTTCTTCTGACAAATGGAGATTCCCTACAGATAACTTAAAAGAAC
AGCTAAGTCACACTCATATGCAAGAATTTACCAAGGCCTAGGAAAAGGGGGGGGGGTACT
GCTTATTATGATAAAGGTCTGCTAGAGCAGAACCCCTGGTGTAGCTTTTACAAGGTTCA
AAGGTTGTAGCATAAAATTGTGACTAGAGTGAGTAACCTTTACCTGTATTAGTCACTCTA
GGCAGAGCTGTTTTATCTTTACTGTAAACATTACCTGGTTCTGTCACTCCTTTGAAGGCAT
TCTCTGTTTTGTGACAGATACTTCTATGTACCTCGCCTGCTGTGACACCCCTACTCCTTTGT
TTCTGTATTATATAAGCCTGGTGTTCCTTTGTGAAAAATTACATCCAGATACAGCACTC
CCTTGTGTCTGTGTCTTTTGTCAATTTCTGGCCAACTCCATGCCACCTGCCAGAACCCCTA
GCTTTTCCACAGATTGAGGGAGGGCGAGCTGAGCCTGGTCCATGGCATCTAAACCACTGTCA
GCTCACTGTTGGTGACTACCTCAAGGTACAAGTCCATTACTAATGAAACAAAATTAGAT
AAGTGTGGGTCCAGGAAGCAGGTTGTACACCTGTCTGAATGAACATTATGAAATGACTG
AAATAAGTTAAACCATCTCTTCTCGTTTGCTAATATAGCAAATAAACCGAGTTTCTGAGC
TGCTGCTGGTGTGTCTCCATCAGAGGGCAGAGCCAGTCTGATCCTAGCTTTCTGTATGTG
TGTCATTGTTTCTTCAGTTCCTGTTGCCCCATTAGGAAATCCTAAGCCATGAAAGCCATG
AATCTGGGAATGACTTTTCTAAGAAATGCCACGTGAACCTTGCCTTTCAACGTTTTGCCTG
TAAACAGATATATGTTGCGCAGTTTATAATCAATAAAGCTTTGAAATAATATATAATC
CATTTCTATTCTGCTTCCACGCTGAGCATCTGTTTCCCCAGGACCACAAGAGCATTTGA
AAAGTAGTGATTATGACCTGCTTTGTTCTGTTACTATAAAAGCTTCATGAAAGGGCAGCC
ATGTTGAAACATGGAACCTTGGGGTGACCTGTATCTGTGTTCTGGGTCTGTCTCACTCATA
TTTGTCTCCAGAATAAATGAGTTTATCAACTTCGAGGAAAAAGTTGTGTGTTTGTATAGCA
CGCCCGTGGAGTCCCAACCTTACTTCTCTGTAATCTGTATATGTTAGAAAAAGTTAATTT
ATGTGATTTCTTCCAACTCCAAATATTTCAAATCTTTTAGCCCTCAGCCTGGGATTTCTTTG
ACTAAGTCTATTGATTGGAAGATCTCAGTGGTTAGGATTTGCAGTCATGATGTTTCATACG
TCAGGCTAAAGCTGAAAAATATGACAAATGAAATGTCAAATGTCAATGTGCTGCGCTGGGAATGTG
AGTGTTAGGGGGTTTTAAAGAAACAAATACCTACTCTAAATAGTTAATAAGTCCCATGGTT
CTATTCTAGTTTTGAATAAATGTTCCCTAGTATACAGCAATTAATTTGAAATGAATAGCTTC
TTATCTTGACCAATCTCAGTGACTTCATCCGTCCCAAGTCATGTTTTTCATATTTCATAAGGAT
AGGTCTCATCAACCACTGTTTATCATTTGGGATCTGCATTTTTCTGTATGCAAAATGATTT
ATTTCTCAGAGCACTGGAATTTGGGTTGAATCATCTTATACGGGCCAAAATCAATGTCTTC
TGCTCTAAACAGAGAGTTACAAGACCTTTTTATGTGGATGGCAGCATTTTAGTCATCCTTA
TGACAGAATGTCAAGTGGAGCTCCCACTGGGGGAGGGGCTGGTCTTTGGCAGGATTCTC
TTGGAACATCACACAAAGAAATTCCAAATTATGAAATGCACATGATCCATCCAGAATGTG
ACTTTTGACTCTTGAACTATGAGCTTTTAAAGTACGTTTGGCTGTTTCAGACCTTGACTTTGAG
GTGAAGGAAAGCTCGCCAACTCCTTTTTATATGTAAACAAATATATCAAGATCTAATGTGA
GACAGATATGCCAGTCCCAAGATCTGTCAATATGACTGAAGACACATTTGCCATGTTATCACT
AAGGCAGGAGAAGGCAAGCTACAGTGAAGCCAGTTCACTATAAAGCTTTATGAGAAAT
AGATAAGAAGGGTTTCTAATTTTTAAATTTTTTTTATTAGATATTTCTTCATTACATTTCA
AAATGTATATCCCAAAAGTCCCTATACTCCCCCTCACCTGTCTCCCTACCCACCCACTC
CCATTTCTTGGCCCTGGTTACTTGTGATAGTGGTCAATGATCCACCAAGCTTTTACATGCTC
ACTATCTGGTCTATTGCAAGAATGGCTGCCGAGCTGATGCACTCAGATACAGACACCTAC
AGCCAAACAGTGGGAAGGAACCTTGGGGACTCTTATGGAAGAAAAGGAGGAAGGGTTATAGG

CCCCGGATGGGGAAAGGAACTCCACAGGAAGACCAACATACTTGGTCAACTAACCTGG
ACCCTTGGGGCTCTCAGAGTCTGAACCAACCAACCATAGAACATTTCATGGGCTGTACCCAG
GCCTCTCCACTCATATGTAACAGATATGTGGCTTGGCCTTCATCTGGGTCTGAACAACTA
GATGGGGGTTAGGGGTGGGGATGGGGGTTATCTCAAAAGCTGTTGCCTGTATGTGGGATA
TGTTCTTACTGAGCTGCCTAGTCTGGCCTCAGTGGGAGAGGAAGCACCTAGCTTTAGCCTT
GTAAAGACTTGAAGTCTGAGGTGTGGGTGGAGGGTATACTCAGGGAGGCCCTCAOCTGC
TAGGAAGAGAAGAGGAGGGGGAAAGACTTGGGGGAGGGGGCAGTGAGCAGGTTGGTAAA
ATGAATAAGAAAAAAAAAATAAAAAATAATTAAGAAAAAAAAAGAAATGGCTGCTGAGCCC
TACTCTAAAAACCATTGCATCCCCCCCCCAATCATTTCAGTGACTACGAATTAATTCATT
GATACTAACAATAGATGTAGGAAACTATTGTTAACTTCTTTGTGAECAGAGTGGTATTG
GAACCTTTTTTATTGAAGCTTTCACACAGAGCCTTGTTCTTTCATTTCCCTGTACATGCATG
TAGCTTAATGATGTTTCAGTGAATTAATAATAACAATGAAGAATAAAGACAACCTGTATTTT
AAGGATTCTTCGTATATATTTAAAAATCTAAGGTGGTCACTGGAAGAAATGTCTTCAGTT
TTTCTATATATGTTTACTCTATCGTATGTTAATTAAATTATATGCAATAATTCATAAAATCTA
CAACATAGTATGTAACCTATAAGAAAGTAAAAACATTCATGAAATTGTGAAGGTTACTTTTC
CTTACCCTCAGAAACACTGGGTTTGAATAATTCTTATTTTGGTATCAGTGAAGAATTGAA
AGAATGTAATAACCTACTAAGGCAAAACATAGAAGTTGAAATTAAAAAGAGTAGACAGGA
GAAGTAATAAGGCAAAATAATGAATATTGCTTTAAATAGTTCTTAATGTATCATCTAACTA
GGGTGTGATTCTCCAGACTTGACTCCATCCAAAATATCCAAAATGACTCTAACCACAGTCA
TTGAAACAATGTGTTGAAAATAATAAACATTTCTACTTGAAAATTTCAGATTTCTCCTACT
TTGCTTTTTATTGCTGTGATAAGCACCATGACCAAAGCAGCTTAT

(SEQ ID NO:18)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

72/85

Figure 49

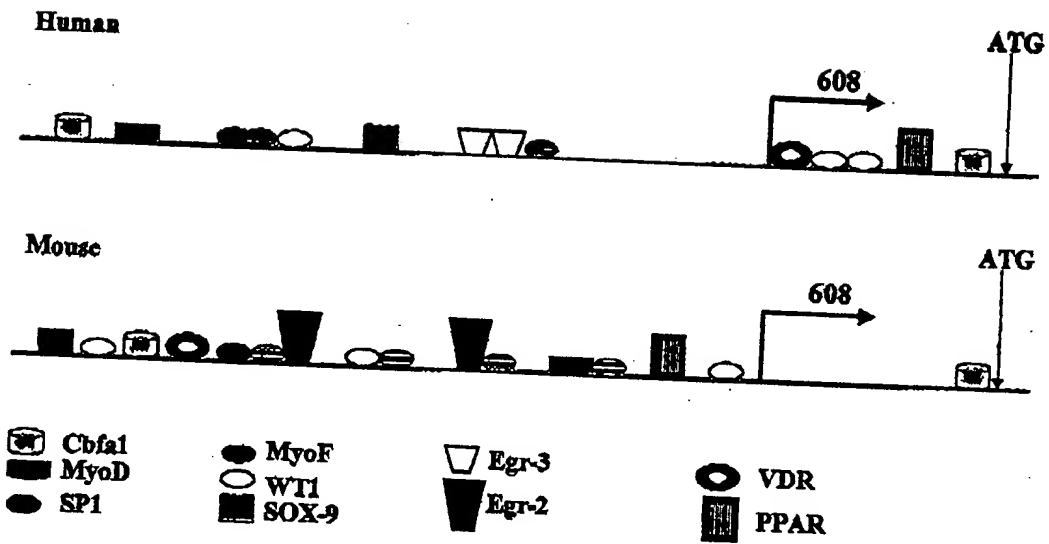


Figure 50

TAAGCCTTTTGGCTCCCGTTGGAAGCAAAGAACGTTTCCTTCAATCAGGTGAAGGCTCT
CCTCAGAAGATTTTCATGTCTCAGCTTATGTTACAAGAGGATTCAAAAGCAAGACAGA
AGAGCTCAGGTATAGCCAACTCTTTTGTAAATACAGTATGAGGCTTAAGTGTACGG
CAACTCATGTGGTATTCATTTGCGGCTCTCTTCTCTTATAACTAACTCTTAAGGTGCA
TATAGTCTCTTCTGTTTCCCAGCTACCTTGCACCATCTTTGTTTATCTAATAATAGCA
AGCTCATCTGCTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTTCAGTGATGTA
ACAGTGACAGTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATATGGGTAAACT
CATTCATAACAGCTCCAGGTTGG (SEQ ID NO:19)

TTTGGAAACCAACCCAGATGCCCTCAACAGAGAAATGGGCCAGAAAATGTGGTCCA
TTTATCCAATGGAATACTACTCAACTTATTA AAAACAACGACTTTCATAAAATTTT
GGCAAATGNATGGTCTGNAGGATCTTGAGTGAGGTAACCCAATCACAAAAGAACAC
TCATGGTATGCACTCACTGATAAGTGGCTATTTGTCTATGGAGTGATTTAAAAGGGA
AGAAGACACATAGCTTTTTGTGTGTATAATATTAAGATGGAAATTTGCCAGTGCTGT
TTGGCTTATGAGTGAATCTTGTTTCAGTGGATTACCGGAAGAAAATAATAAGTGAAC
TGTAGGAAGAAGTAGTTAATCAAGGTGACAAAGTATCCTGACACATTGGGAAAAGA
CCACAGTCCAGGAACTGAGTCTTAAGGATTCATATTA ACTCCAGTTCCCCATGTGC
AGCTCTGAGACTTTGGCAGATCAGACACTTA ACTTCACCAGCTTCCTACACAGAGCA
GTTACTATCCTTGCCTTCACACATGGAGTGTGCCATTAAGTGCCTGAACATGAGTCT
GACTTGTTAATAATCTTTAAAATCCAATTGTGTGTAAAGTATGTGACCAAAGAGCAT
GGTCATGCTATTAACCTTTGATGTTCTATGGACTCTTAATTTTATGGTAGAAATGTCA
ACAAGCTTGTGGAGGCTGGAAGATACAAGGCTTAAGAGGATGGCCTTTCAGTTTTG
AAAGTAATTCAGTATGTGTTCTGGCATCCCTTTTCTAAAGCAATTTAACCCCCCAA
GTAGGCATAATTTTAATGCTTACTTCATCAGAATATATCTAATTGACTCTTCTAAAAA
GACTTTGGTATGCATAGGATCTAAATGTAAATGTGATTTACTGACATAATAAATAGG
AGAACTGAGCTAGAATAGGTATAAAATATGTGCTGGCTTTCTAATAGGTCTTATAG
GTTATATAAGAGGTGGGAAAGGAATATTTGAAACATCTAGAAGTAAAATGATCCTG
AGTAGCGATCCTGGGAAAATACGTA CTCTAACACACTGCAATCATCTCTCTGTGGTT
TGCTGGAGCTGAGGTCTGGAAGGCTCGACCTTGGTTAGAAATAACCTACCGAATAC
AGAGCTATGACGTTAGTCTGGAAGGAGCTTTGGAAGAATGACAAGCTGTAGCTGCC
CAGAACATACTAGATGCCATATTTCCAAGGCAAGTGTCCACATGCGGACATCTTAAG
AATATGGTTGTCTCTGCAGTGCTAAGGACCTTGTTTCGTGCCACACAGGTCTCCAGGG
T TAGTGCTAACTCTGACTGCTTGACTCTTTAATTCTCCCTTGATCATTAAATGACTAGA
AATCACTTGGTGATTAGCAACTGGATATGGAATATTACTTAATTTGTACCCAAGCCA
GGCCACCTCAGCTTTGGCAGCTCCATTCTGTGGAGCCCAGTCACGTGGGTTTG
AATCAACTGTACTGTTTCTACTTACAAGACGCATTACCTGAGATGAGTCATTTTTCTT
CACAAGTCTTTTTAGAAAGAGTCAATTAGACATATTCTGATGAAGTAAGCATATAAAG

TGAGAGCAGCATGAATGTGTTCCATGTATGCTCATGGATGCTATTATAATGTGGAAA
 TAAACTGACTTTAAAAAAAAAAGCTTATGATACTTGTACAGAGTAAATCTTCCATA
 AATATCATCTGCATTTATAAATTATTTTCATAATCCATCAATTA AAAAACCTTTAGAAA
 TTTTGTTAACACAAAGATCCCTAGGCCCTGCCCTAGGATGGTCTGTATGGTGGGCC
 TGAGAGATGGAGCTTAAGA ACTTACTTGCTCCAGGAGCACATCTTCAGAACATCTGC
 CTCAAAACATTTATCCCAAATGCTCATCAAAGGCTCACTCACATGTGCTTCAACCAC
 AGGGATTAAACAGTCATTTTAGTCACATTTCTCAAACGGTGGAAGCCTGCTAGAGGA
 ACAGGATGTATCAGGATAACATCCAACCTTACAAAAGGATGTCATAACCCTCACCA
 CAACAAACAACAACGACAACAAACCCATAAAAATTATCACGGCAAATGAACTAAGC
 CATATGCAGAAAAAGTATTATATGTTCTCATTGTGGGGTGT TTTTCCTTAATAGTCAA
 ATATGCAGAATATAGACAAAGATGGTTTATGCAAGTGGGGATGGCGAAGGATACTT
 GTAGATTAGAGGACACAAAGCAACA ACTACAGAGTGAAGTAATCCAGAGACTTAAT
 GTATAATATGAGGACTGTATTTAATAATTCTATTTAAGATACACAGCAAACGAGTGT
 ATCTTACTAACACACACACTTACATAGAGAGAATAAAGTGATAGATACGTTTGT TTT
 ATCTTCATGTAGCTGATAATTTTCATATTGTACACCTCAAACATAGATAACCAACAAA
 GAGGAAGAGGATAGGTGCCTCTCCCAGGGCGGAAGAGTACATTTCGAAAGTCAGACA
 CCATTGTGTAGATGTACCACATGGAGGAGCTAGAGAAAGTAGCCAAGGAGCTAAAG
 GGATCTGCAACCCTATAGGTGGAACAACATTATGAGCTAACCAGTACCCCGGAGCT
 CTTGACTCTAGCTGCATATATATCAAAAAGATGGCCTAATCGGCCATCACTGGAAAGA
 GAGGCCATTGGACTTGCAA ACTTTATATGCCCCAGTACAGGGGAATACCAGGGCCA
 AAAAGGGGGAGTGGGTGGGCAGGGGAGTGGGGGTGGGTGGATATGGGGGACTTTT
 GGTATAGCATTGGAAATGTAAATGAGTTAAATACCTAATAAAAAATGGAAAAAAA
 AAAAAAAAAAAAAAAAAAAGGAAGGTCAGACACCTCACTTCACTGCTATCTCAACTTG
 CAAACAGAAGGGGAGTCACAAACCCAGGACAAACCACAGTGATTGAAGCGTCTTTG
 AATGTTATTGCTGTTGTTGTTACCACCATCATTAGCATATATTCATTGTGAAA ACTTA
 CGGGGTCTATGACATGTTTTTTTATTCAAGTATATCACATGCTGTCAGCATATTTGGC
 ACCACTACCAGCCCCAGCCCCCTTTGCCCCGCCCCAACACACACACACACACACAC
 ACACACACACACACACACACACACACACACACACACCTTTACCTTCTCCTGGGCA
 TCATCTGCTCACTCACCCACCCAAGCTTAATCCTTTTCTTCCCTGCAATAGTACCTC
 TCCTATTTTTTATGTCTAGGTTCCCCCTCCCCCTGTTAGGAGATGGGAGAGGTCACGA
 AAGAAAGGAATTTGTAGCCCTTGAGCCAGCCCGGGCCACAGAGCCTGCCACCAGAC
 AGGAAAAGCCCAGGGCTTACCAGCACAGGAGGAGCAA ACTCGCAGGCGAGCCTGG
 GTTGGCGCTGGTGGTCCCCGGGTGCATGGCCCCGCCATTCCCAGAAGCCGAGGCTATA
 GCTGCGTCACCTGCCCCGCCCTCCTCCCGAGTGAAGACCCCTAGAGGCTGAGCAGAC
 CCCAAAGGCGGTGCAATTCCATTGGCCCAAGGCAGAGGTGAGCGGCTGCTAATCCC
 CTCGGGAAGTGAAGGGACCCAGAGAGTCTGGTAGATGTGGGAGCTGGGGTT CAGGG
 CGAGACAGAGGGTGGGATGGGCAGAAGGGTCCAGGAAAGGAAAGTACTGGAGGGG
 AGTTGGGACAAAAGCAGCGACCAAGGGAACATCGCTTCAGTGACTGAAGCCAGGCA
 AAAGGAGCGGGAAGGATTATATGTAGCCTGGGACGCTTTCATAAACACTGATGACG
 TGTTTGTGCAAAGCAAGCAATTTGAGGAGAAACGCCTGGGACGTCGGAAAGAAGGA

GTGATCGATTAGTACTTGTAAGTTTAGGTGAGTTTGAGAACTAACTAACCTATACTA
TTGAGGGAGAAGGAAGAGCATTCCAGCAGCAGCAGCAGCAGCAATCAGATAA
AGGAAAGCTTTGGTTAGTTTGGAAATGTATGATACCATTAATAAACAGAAGCGCCT
CCAGTTCTCTGAAGAGTCAGTCCCCCAGCTAGTGAAGACTAAGCCTACTAAGCCTTT
TGCTCCCGTTGGAAGCAAAGAACGTTCCCTCAATCAGGTGAAGGCTCTCCTCAGAAG
ATTCCTGTCTCTGCTTATGTTACAAGAGGATTCAAAAAGCAAGACAGAAGAGCTCAG
GTATTGCCAACTCTTTTGTAAATACAGTTTGAGGCTTAAGTGTACGGGAACTCATG
TGGTATTCATTTACGGCTCTCTTCTCTTATACTAACTCTTAAGGTGCATATAGTCTC
TTCTGTTTCCCAGCTACCTTGTACCATCTTTGTTTATCTAATAATAGCAAGCTCATCT
GCTTTTTTAATCATCACGCAGAGAGTATTCAAAAATATTCAGTGATGTAACAGTGACA
GTGTAGGCATAGAAGTAATCATTAGTAAATCTTAATTTGGGTAAACTCATTCATAA
CAGCTCCAGGTTGGGAGGGATCACTGAGCCTTCGCCACGTGCGGGTTAAAGATATTT
TCTAACAAGAGAAGCAGAATTCTTCCTTGCCATGCTCCCCATCACTGTGTACGTAA
GCAGAGGGGTGTTTCCAAGCAGAGAAAGAGCAGACAGTGTTATGCCTGCAAAGTCA
GAGACTCAGCCCTCCCAGCTGGTCAGTTTACTGTCCTCCCGGTCATTAGTTGGCTCTG
AAAAGGCCCATGTGTCCTTATTGGCAAGGACTTGCAGACATGCTAGAAAGAAATTT
GACCTTTTTTTCTAGTGGGTATTACAGCTGTAAAAGTATTTTGGAAAGGTAAAGCCA
AATAAATAAAACACATATTAATAATACAATGTTACAAAAATTGATCATATAAAGA
AGTACATTCATAAATGCAATGTGAAAAATATATATAATTTTTATCTATTTACTGGTGC
AAAGTTTTCTAAATTGCACATGTACTATTTTTATATTTATAAAAAATATTTTTAAAATG
TATATAAAAGTGTAAGGCTCTTGGTCAAACAAGAGAGTTAAATTTACAACTTTA
ATTGTCCCGATAACATTATTATGATCTCTAATGACAGGGATCCTGCTTTTCATTGGGA
AATGAGAAGCTATGAAGATATGTTTACAATAATAAGCCCATTTAGTGATAAAGTCCA
ATGGGAAGCTAGCACACACTGGTTTATAAAGAGAACAGTTTCCTGAGTCTATGCAA
GTTTACACTCTAGGGAATAAGAGTTCCTCTTTCTCCAGATTTCACTAGCATTGTGT
CATCATTTATCTTCTTGATGATGAGCATTATAAGTGGAATAAGATAGGATCTCAAAG
GAATGTCAATTTGGATGCCCTGAACAATCTTTCAGGTCTTTCTTTCAGTTCCTAGTC
TATTCATTTATTGGATAATTGGGGGGATGGTGGTAATTTTTTTGCAGTTCTTATGGAA
TTCCAAAAAACAAAAACAAACCAACCAACCAAAAAACCTCTGAACTAGAACTACC
AATCCATTACTGGGTATGTAACAAAGAGAAATCTGCACAGAATTTATTGCTACATTG
TTCATTATTCACGACAGCCAAGAATGTGGAACCAACTTACGTAGCCGTCAAAATATG
AACGGATAAAGAAAATGTGGAAATGTGTACAACAGAGTCCCATGTGGCCATAAAAG
AGTGAAATCATGACATATGCAGGAAATGGATGCAACTGGAAATCAATTGGGCTAAT
CAAAACAAGACAGACTCAAAAAGGAAACACCGTGTAGCTTCTCTGACAAACAGAAG
CTAGATTTACACTTGTACGTGCGCATGTGTGTTTAGAATTTTATTTAGTTATACACTA
TTCTAATCTGTGAGTGTGTATAAAGGCATGCATGTAAAGCAAAAACAAGCTAGCTG
GGGTGGGTAGGAGAGAAAGCAATGAGAGGAGTTAATAAGAACGAAGCATAGTAAC
ATAGGTGCCAGGATGAAATGCATTAATTTGTATGCTAACTAAACCACAGACAGGAG
GCACACGTTCAAACCAGGGTGAAATCCCAGCACAGAGAAGGGGAAGTAGACACAA
AGTTTCGCCACTAACCAAGAAGCCATTTGCAGTTGCTGCCTGCTGGGAAGGGGCGTT

CCAGTTTTCTCCAGTCTGACACTGTGTATAACAACCAGTTGACAATACAAAGTTGGC
ATGATGGATGGTTTTTGTGCTATTTTTCATTTTTTTCTTACTGTTTTGTTGTTGTGGT
GGTTGTTGTGGTGGTGGCTGTGGTTTTTCATTTGTTTCTTTTGAGAGAGAGAAGGAAC
ATGAAATTGGGTGGGTAGGAAGCTGGAAACGATCTGGAAGAAGTTGGGGAAAGAG
AAAAATTGTATGGAGCATATTTAAACAAACAAACAAACAAACAAAGGTTTCATTTT
GCCACAAAAAGGTGTGAATTAAATTAACCAGTTACGACTCTTAAAGAAAATATTCCC
AATTATTCCCAGAGTTGCTATGTATGCTGTGCCTAGGACTTTGCTTGAAGTGGCCCTA
TAACTCTGGTGTGGTGTCTTTTCAGGATGCAGAAGAGAGGCAGGGAAGTCAGCTGCT
TGCTGATCTCCCTCACTGCCATCTGCCTGGTGGTCACCCCTGGGAGCAGGGTCTGTC
CTCGCCGATGTGCCTGCTATGTGCCACAGAGGTGCACTGTACATTTTCGGGACCTGA
CCTCCATCCCAGACGGGCATCCCAGCCAATGTGGAACGAGTCAATTTAGGGTGTGTG
GACCTTGCTGATCTCCTTCTCAGAGAGGGACCACTGATTTTCCTGGTACTTTGCCCC
CCAAACACCTGTGATTACTTTTAATAGTTTTCTTCTAAAATGGGTTTCATACAAACCTT
ATATTGTGGAGACAATGAACATTTTATCCCAATAGTCTTTTACTAGAAGTGAAGCC
CCTCTTAGTTGTTTGGGAGCCTCATAATTATGGGGCAGCTTTATTCTGAATGAATTTT
AAATGAAAAAGATACAGTTTCTGTTAACAATCATTATGATACCAAGGAAGAGGAAT
TGTCATTGAATATTTTAAAAAAGCATTTCTTTTGCAATTTATAAATACCCATTACAAA
ATGGCTTACTTAAAATACTTGCCTTACTAAATCTGACAAATTATGGTGATATTTTGAA
GGTTTATGAAAATTTGTTTATGTGTATAAATGCACAAGAAATGGGATATGCCATCAC
CTATGTGCCATTAGTGAGCATGTACAGTATGCCAAACACTATTGTTACGTTTGGAG
GAAGTAATGGGGGTGGGGGAGCAACAAGGGTTATAACCGTATACCCAGTGCCTTGG
AAGCGATTGCAAACAGTAAAGACTGACATTGTGTTCTCCCTATGAGGGAGGGGCCTT
GGGCTGAGCACTTTGCAATGAGCATTGTGCTCATTGTGCTGGCAGGTTTTATGATAAC
TTGACCCAAGCTAGAGTCACTGGAGAGGAAGGAAGTCAACTGAGAACATGCCTGA
AGAAGATCAGATTATAGGCAGGCCTGTGGGGCATTTCCTTAATTAGTGATTTCATGGG
GCAGGGCCCAGTCCATTGTTTCGTGGTACCATTTCAGGCACTATTAACAAAAA
AACAGGCTGAGCAAGTGTCAGGAGCAAGTCAGTGAGCAGCAGCCCTAATGATCTC
TGCATCAGCTCCTGCCTCCAGGTTCCCTACCCTATTTGAGTTCCTGTCTAGCTCCCTA
CAGTGATGAACAATGATGTGGAAGTATAAGCCAAATAAATCCTTTCTTCCCCAACTT
GCTGTTGGTCATGATGTTTCATCACAGTGATAATAGTCCTCATGAAGATGCTGGTGT
TTATAACACCTTTGGACTAAATTCTGTTATCTATAGCTGAGGAAAATGGAGCATAGA
AAGTCTCCAGACTACACCAGAGTGTAATCTGGGCCTGAGCTTAGAATCACACCCAC
GTGCACTCCACTGCCGGGGCTTCTTAACCGGAACACAGTTGTAAAAGGGAATTTTCT
GTTTGTTCATTTTGACATGTGGACTTTAATTGACGATTCATCTGAAGCTGAAAATG
ATTTTTTTCCAGGTATAACAGCCTCACTAGATTGACAGAAAATGACTTTTCTGGCCT
GAGCAGACTGGAGTTACTCATGCTGCACAGCAATGGCATTACAGAGTCAGTGACA
AGACCTTCTCGGGCTTGCAGTCCTTGCAGGTGAGATAGGTAGAGGGTGTGGAGGC
TGAGAAGAGAGGTGCAACTGTGGGTATACCCAAAAGCTGCTGATTCCCGTGGGAG
ACATTCTATAAGCATTCTATAAACTAGAGGCAGATATCAAGGAAGGATTTCAATTGT
AATGCAATTTTATGAGAAAATTTGAATATTAAGAAAATGCTGGGGAAAATGCTTAC

ACAATTGCGAGGACCTAATTTAGGATCTCCAATAGCCACATAAAAAGCACAGCATG
 GCGGCAGACACCTGCAATTCCTGTCCCTGGAAGCACCTGTTTCAGAATCCCAGAGACT
 CATTGGCCAAACACTCTATTCAATCAATGAAGTCCATATTCAGTGACAAAACCTTGAC
 TCAGAAACTAATGTGGAAAGCATCAGGAAGACAGCCAACATCTGGTCTCTACTCAT
 GCATGAATAAGGGATCCCAGAGAGAAGGGAAGGAAAAAGGAAGGAAGGAAGGAAG
 GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAGGGAGGAAAGGAGGG
 AGGGAAGGAAGGAAGGGAAGGGAAGGAAAAAGAGATGGGGAGGGAGGGAAGGAAG
 GAAAGGGGGAGAAAGAAGAGAAGAAAGGAAAAATAAATAAATTTTCAGGGATTATT
 ACACCTTTAAATTTTATCCATAAAAGGTCATTTCCACCTGTTTGTCTGGAAGTAGAGT
 GGGATCCCTTATATAAGGGCAGTCTTTAACATAGTAGCATTTTATAAACCATTACAA
 ATTTTGAGTTTTCTCTACTTTTTATCCTCTACCATCTTCAAACCTGAAACTACAATTATT
 CCCACAAATGAAGAAAATGCTGTAAAGAGTTTTTCACACACCGAAGTGGAAGAACTTAA
 GGATTAGACAAGTCTAACAATGAGAATGGGGAGAACAAAAAGAGACTGCACAGGG
 AGCCCTTTCTCTGCTTATAATCTTGACACTTGAGAAGCTAATTGACGCTGCATGACTA
 CTCAACTCTTTAAGCAAACAATGCTGTTGTTTCATGAAAAGCACATAAAGTACATAT
 GTCCCATAATATTCATCAAAATTTGCATGCAGCACATAATAGCAATCAAAGCAATAA
 CACCCACTGTTTCACAGAGACTTTAAACATGAAACTGGAACATATGTCTAGTGTTTTGA
 CTTAGGGTACATAGTATGCTGTGTCTGTATGTACCAATGTTGATTTAGGTCATCAGA
 CAGCATTTGGAACATGTATCTTCAGGAGGAATCATTATGTATCCTGCATGAAATTC
 TCCACCTATGTTTATTCTCTTAGCCAGGTTTTTCTCTGATGGAGAAACATTGGGTTTTG
 AGGTTTTACTCCCAGGTAACATTTAGGGAAAAGCTGTCTATGTTCTCAGTTTGGCTTT
 TATTTATGAGGGATGTTGGTATTCCAGAAAATTCTCTTTTGAAGAGATTACAATTTA
 GGTCAAAACAGAAAAATATGTAAAAAGTTATTGTTTTATTAGTATTTTCATGTTCTTT
 TCTTTTTTAAAAATGGTATGCTTAGAATAATTAAGATTAGATTAGATTAGATTAGA
 AAATAATCAGAGAGGGATTTGATGAATGCTAAAGCATCATGAAAAATTCAAAATTT
 TTTGCTTCTAATTCAGAATCAATTAAATTCATATTACTATAAAAGACAGCACGCCAG
 ATGTGTGCCAGCTGAGGAGTGGATAAACTGTGTAACGTGAGTGCTATGTAGAAACA
 GAAAGGAGTGAAGGGTTGATGTGCGCTGCAACATCTTGAAAACATTCGGCTACATG
 ATGGAAGCCAGGCACAAAAAGCCACATATTGCATGGTTATGTTTATATGAAATGTTT
 AAAATACATGGATTCTTAGCAAACAGAGTAAGATGTTACTTAGGGTCAGGAAAAGA
 TTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA
 TCCTAAGACGAAATAGTTGAGGTAGATATAGTTATATCCCTGTGGATATTGTAATAA
 ACCAGCATGCTGTGCTCTGAGAAGGGCCTAATGAAGGGGCAGGAGGAAGTGAAATG
 AGATGGTAGAAAGGAAAGTCATATACCATGGCTTCTCTCGTGGGTGGAATCTAGAT
 ATGTTAATATATTGACATAAAGGAAGGAATTGTTTAGGGAAGGATCAAAACCAACA
 GGAGTGAGGGAGACAATAGGAACCAATGAGAGGCAAAGTTCATGGTCAATGTGTGT
 GGAGACACCATAATAAAACTCCTTTTTTGTGTTGCTAACTAAAACCACTAAAATCTAA
 AAACAAAACATTTTTGCACAAGAATTATTTATTATTCAATAAAGATGTTTAAATGGG
 GGAAGTTGAAGTTCATTGATAGTCTCATAAATCTTAAATGTATTTAAACTGCTTTTTA
 CGTTTTTTATTATTAATTACTCTTGCTGTCATTATTATCATCATCATTATCGTCATCAT

CATCACTAATGCTTTTCACCATACACAAATGTAGGCAGAAGAGTGTAATCCACTTAG
TGAGGCAATCTTGGAGAGGGAAAGGAAGCGGATGCGGGGCAGAGGCACACAGGAG
GACAGTGAGAGGGAAATGAACAAGAAAAAATGTGGACACATGCACAAAAATTCCA
TAGTCCACTACATTACTTTGTATTCTAATATTAAGAAAATAATAAACCCATTTCTGTG
CACTTATCACCCAGGCTCAACAGTTATCTTGGCCACAGATCCTGTCTCACTGCATCCT
GTCCACCTGAGTCCACTTAGCGTTCTGAATCCAATCCAGGGCATGATGCTTACTCCT
ACACAGAACTAAAGATTAAAGAGAGTTTAAAAGTAACCATGACATCTCTCTGTTCT
TTAGCGATAAGTTCTTAATATTTATGGCTGCTTGTGTATGTTCTAATTTCTCTAATATT
GTCACATTTAGTTGGCAACTACTTTGTTTGAATTGAGTTGGAGTTAAGGTCCCATAG
GATTAATCTCAACATATTTCTATATTTATAAACTTTTCTCTCTTTGTGAAAGTTCCTTT
GAGAAAACAAATATGCCCATATCTTTCTTTACAGGTCTTAAAAATGAGCTATAACAA
AGTCCAAATAATTGAGAAGGATACTTTGTATGGACTCAGGAGCTTGACCCGGTTGCA
CCTGGATCACAAACATTGAGTTTATCAACCCCGAGGCGTTTTACGGACTCACCTT
GCTCCGCTTGGTACATCTAGAAGGAAACCGGCTGACAAAGCTCCATCCAGACACAT
TTGTCTCTTTGAGCTATCTCCAGATATTTAAAACCTCCTTCATTAAGNACCTGTA
GTATGATAACTTCATTGACCTCCCTCCCAAAGAAATGGTCTCCTCTATGCCAAACC
TAGAAAGCCTTTACTTGATGGAACCCATGGACCTGTGACTGCCATTTAAAGTGGT
TGTCCGAGTGGATGCAGGGAAACCCAGGTAACATCTTGTTTGTGTTTCTTTTTTT
ATARKACGTATTTTCTCAATTTTCAATTTAGAAATGATATCCCAAAGTCCCCCATAACC
TCCCCCCCCACTTCCCTACCTACCCATTCCCATTTTTTGGCCCTGGCATTCCCCTGTACT
GGGGCATATAAAGTTTGCCTGTCCAATGGACCTCTCTTCCAGTGATGGCCAACTAG
GCCATCTTTTGATACATATGCAGCTAGAGTCAAGAGCTCTGGGGTACTGGTTAGTTC
ATAATGTTGTTGCACCTACAGGGTTGAA (SEQ ID NO:20)

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100

Figure 51

MPKRAHWGALS VV LILL WGHPR VALACPHPCACYVPSEVHCTFRSLASVPAGIARHVE
 RINLGFNSIQALSETSFAGLTKELELLMIHGNEIPSIPD GALRDLSSLQVFKFSYNKLRVITG
 QTLQGLSNLMRLHIDHNKIEFIHPQAFNGLTSLRLLHLEGNLLHQLHPSTFSTFTFLDYFR
 LSTIRHLYLAENMVRTL PASMLRNMPLEENLYLQGNPWTCDCEMRWFLEWDAKSRGIL
 KCKKDKA YEGGQLCAMCFSPKKLYKHEIHKLKDMTCLKPSIESPLRQNRSRSEIEEQEQ
 EEDGGSQ LILEKFQLPQWSISLNM TDEHGNMVNLVCDIKKPM DVYKIHLNQTDPPDIDIN
 ATVALDFECPMTRENYEKLWKLIA YYSEVPVKLHRELMLS KDP RVSYQYRQDADEEAL
 YYTG VRAQILAEP EWVMQPSIDIQLNRRQSTAKKVLLSYYTQYSQTISTKDRQARGRS
 WVMIEPSGA VQRDQTVLEGGPCQLSCNVKASESPSIFWVLPDGSILKAPMDDPD SKFSIL
 SSGWLRIKSMEPSDSGLYQCIAQVRDEMDRMVYRVLVQSPSTQPAEKDVTIGKNPGES
 VTLPCNALAIPEAHL SWILPNRRIINDLANTSHVYMLPNGTLSIPKVQVSDSGYYRCVAV
 NQQGADHFTV GITVTKKGSGLPSKRGRPGAKALSRVREDIVEDEGGSGMGDEENTSR
 RLLHPKDQEVFLKTKDDAINGDKKAKKGRRKLKLWKHSEKEPETNVAEGRRVFESRRR
 INMANKQINPERWADILAKVRGKNLPKGTEVPPLIKTTSPPSLSLEVTPPFPAVSPPSASP
 VQTVTS AEESSADVPLL GEEHVLTGISSASMGLEHNHNGVILVEPEVTSTPLEEVVDDL
 SEKTEEITSTEGDLKGTAAPT LISEPYEPSPTLHTLDTVYEKP THEETATEGWSAADVGSS
 PEPTSSEYEPPLDAVSLAESEPMQYFDPDLET KSQPDDEDKMKEDTFAHLTPTPTIWNDS
 STSQLFEDSTIGEPGVPGQSHLQGLTDNIHLVKSSLSTQD TLLIKKG MKEMSQT LQGGNM
 LEGDP THSRSSSESEGQESKSITLPDSTLGIMSSMSPVKKPAETT VGTLLDKD TTTVTTPR
 QKVAPSS TMSTHPSRRRPNGRRLRP NKFRHRHKQTPPTTFAPSETFSTQPTQAPDIKISS
 QVSSLVPTAWVDNTVNTPKQLEMEKNAEPTSKGT PRRKHGKRPNKHRYTPSTVSSRA
 SGSKPSPSPENKH RNIVTPSSETILLPRTVSLKTEGPYDSLDMTTTRKIYSSYPKVQETLP
 VTYKPTSDGKEIKDDVATNVDKHKSDILVTGESITNAIPTSRSLVSTMGEFKEESSPVGFP
 GTPTWNPSRTAQPGRLQTDIPVTTSGENLTD PPLLKELEDVDFTSEFLSSLTVSTPFHQEE
 AGSSTTLSSIKVEVASSQAETTTLDQDHLETTVA ILLSETRPQNHTPTAARMKEPASSPS
 TILMSLGQTTTTKPALPSPRISQASRDSKENVFLNYVGNPETEATPVNNEG TQHMSGPNE
 LSTPSSDRDAFNLSTKLELEKQVFGSRSLPRGPDSQRQDGRVHASHQLTRVPAKPILPTA
 TVRLPEMSTQSASRYFVTSQSPRHWTNKPEITTYPSGALPENKQFTTPRLSSTTIPLPLHM
 SKPSIPSKFTDRRTDQFNGYSKVFGNNNIPEARNPVGKPPSPRIPHYSNGRLPFFTNKTLSP
 PQLGVTRRPQIPTSPAPVMRERKVIPGSYNRIHSHSTFHLDFGPPAPLLHTPQT TGSPSTN
 LQNIPMV SSTQSSISFITSSVQSSGSFHQSSSKFFAGGPPASKFWSLGEKPQILTKSPQTVSV
 TAETDTVFPCEATGKPKPFVTWTKVSTGALMTPNTRIQRFEVLKNGTLVIRKVQVQDRG
 QYMCTASNLHGLDRMVVLLSVTVQQPQILASHYQDVTVYLGDTIAMECLAKGTPAPQI
 SWIFPDRRVWQTVSPVESRITLHENRTLSIKEASFSDRGVYKCVASNAAGADSLAIRLHV
 AALPPVIHQEKLENISLPPGLSIHIHCTAKAAPLPSVRWVLGDGTQIRPSQFLHGNLFVFP
 NGTLYIRNLAPKDSGRYECVAANLVGSARRTVQLNVQRAAANARITGTS PRRTDVRYG
 GTLKLDCSASGDPWPRILWRLPSKRMIDALFSFDSRIKVFANGTLVVKSVTDK

DAGDYLCVARNKVGDDYVVLKVDVVMKPAKIEHKEENDHKVIFYGGDLKVDCVATGL
PNPEISWSLPDGSLVNSFMQSDDSGGRTKRYVVFNNGTLYFNEVGMREEGDYTCFAEN
QVGKDEMVRVRVKVVTAPATIRNKTYLAVQVPYGDVVTVACEAKGEPMPKVTWLSPTN
KVIPTSSEKYQIYQDGTLLIQKAQRSDSGNYTCLVRNSAGEDRKTVWIHVNVPKING
NPNPITTVREIAAGGSRKLDCKAEGIPTPRVLWAFPEGVVLPAPIYGNRITVHGNGSLDI
RSLRKSDSVQLVCMARNEGGEARLIVQLTVLEPMEKPIFHDPISEKITAMAGHTISLNCS
AAGTPTPSLVWVLPNGTDLQSGQQLQRFYHKADGMLHISGLSSVDAGAYRCVARNAA
GHTERLVSLKVGLKPEANKQYHNLVSIINGETLKLPPGAGQGRFSWTLPNGMHLEG
PQTLGRVSLLDNGTLTVREASVFDRGTYVCRMETEYGPSVTSIPVIVIAYPPIRITSEPTVI
YTRPGNTVKLNCMAMGIPKADITWELPDKSHLKAGVQARLYGNRFLHPQGSLLTIQHAT
QRDAGFYKCMANKILGSDSKTTYIHVF (SEQ ID NO: 21)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171

Figure 52

ATGCCCAAGC GCGCGCACTG GGGGGCCCTC TCCGTGGTGC TGATCCTGCT
 TTGGGGCCAT CCGCGAGTGG CGCTGGCCTG CCCGCATCCT TGTGCCTGCT
 ACGTCCCCAG CGAGGTCCAC TGCACGTTCC GATCCCTGGC TTCCGTGCCC
 GCTGGCATTG CTAGACACGT GGAAAGAATC AATTTGGGGT TTAATAGCAT
 ACAGGCCCTG TCAGAAACCT CATTTGCAGG ACTGACCAAG TTGGAGCTAC
 TTATGATTCA CGGCAATGAG ATCCCAAGCA TCCCGATGG AGCTTTAAGA
 GACCTCAGCT CTCTTCAGGT TTTCAAGTTC AGCTACAACA AGCTGAGAGT
 GATCACAGGA CAGACCCTCC AGGGTCTCTC TAACTTAATG AGGCTGCACA
 TTGACCACAA CAAGATCGAG TTTATCCACC CTCAAGCTTT CAACGGCTTA
 ACGTCTCTGA GGCTACTCCA TTTGGAAGGA AATCTCCTCC ACCAGCTGCA
 CCCCAGCACC TTCTCCACGT TCACATTTTT GGATTATTT AGACTCTCCA
 CCATAAGGCA CCTCTACTTA GCAGAGAACA TGGTTAGAAC TCTTCCTGCC
 AGCATGCTTC GGAACATGCC GCTTCTGGAG AATCTTTACT TGCAGGGAAA
 TCCGTGGACC TGCGATTGTG AGATGAGATG GTTTTTGGAA TGGGATGCAA
 AATCCAGAGG AATTCTGAAG TGTA AAAAGG ACAAAGCTTA TGAAGGCGGT
 CAGTTGTGTG CAATGTGCTT CAGTCCAAAG AAGTTGTACA AACATGAGAT
 ACACAAGCTG AAGGACATGA CTTGTCTGAA GCCTTCAATA GAGTCCCCTC
 TGAGACAGAA CAGGAGCAGG AGTATTGAGG AGGAGCAAGA ACAGGAAGAG
 GATGGTGGCA GCCAGCTCAT CCTGGAGAAA TTCCAAGTGC CCCAGTGGAG
 CATCTCTTTG AATATGACCG ACGAGCACGG GAACATGGTG AACTTGGTCT
 GTGACATCAA GAAACCAATG GATGTGTACA AGATTCACTT GAACCAAACG
 GATCCTCCAG ATATTGACAT AAATGCAACA GTTGCCTTGG ACTTTGAGTG
 TCCAATGACC CGAGAAAAGT ATGAAAAGCT ATGGAAATTG ATAGCATACT
 ACAGTGAAGT TCCCGTGAAG CTACACAGAG AGCTCATGCT CAGCAAAGAC
 CCCAGAGTCA GCTACCAGTA CAGGCAGGAT GCTGATGAGG AAGCTCTTTA
 CTACACAGGT GTGAGAGCCC AGATTCTTGC AGAACCAGAA TGGGTCATGC
 AGCCATCCAT AGATATCCAG CTGAACCGAC GTCAGAGTAC GGCCAAGAAG
 GTGCTACTTT CTTACTACAC CCAGTATTCT CAAACAATAT CCACCAAAGA
 TACAAGGCAG GCTCGGGGCA GAAGCTGGGT AATGATTGAG CCTAGTGGAG
 CTGTGCAAAG AGATCAGACT GTCCTGGAAG GGGGTCCATG CCAGTTGAGC
 TGCAACGTGA AAGCTTCTGA GAGTCCATCT ATCTTCTGGG TGCTTCCAGA
 TGGCTCCATC CTGAAAGCGC CCATGGATGA CCCAGACAGC AAGTTCTCCA
 TTCTCAGCAG TGGCTGGCTG AGGATCAAGT CCATGGAGCC ATCTGACTCA
 GGCTTGTACC AGTGCATTGC TCAAGTGAGG GATGAAATGG ACCGCATGGT
 ATATAGGGTA CTTGTGCAGT CTCCCTCCAC TCAGCCAGCC GAGAAAGACA
 CAGTGACAAT TGGCAAGAAC CCAGGGGAGT CGGTGACATT GCCTTGCAAT
 GCTTTAGCAA TACCCGAAGC CCACCTTAGC TGGATTCTTC CAAACAGAAG
 GATAATTAAT GATTTGGCTA ACACATCACA TGTATACATG TTGCCAAATG
 GAACTCTTTC CATCCCAAAG GTCCAAGTCA GTGATAGTGG

TTA CTACAGA TGTGTGGCTG TCAACCAGCA AGGGGCAGAC CATTTTACGG
 TGGAATCAC AGTGACCAAG AAAGGGTCTG GCTTGCCATC CAAAAGAGGC
 AGACGCCAG GTGCAAAGGC TCTTTCCAGA GTCAGAGAAG ACATCGTGGA
 GGATGAAGGG GGCTCGGGCA TGGGAGATGA AGAGAACACT TCAAGGAGAC
 TTCTGCATCC AAAGGACCAA GAGGTGTTCC TCAAAACAAA GGATGATGCC
 ATCAATGGAG ACAAGAAAGC CAAGAAAGGG AGAAGAAAGC TGAAACTCTG
 GAAGCATTCG GAAAAAGAAC CAGAGACCAA TGTTGCAGAA GGTCGCAGAG
 TGTTTGAATC TAGACGAAGG ATAAACATGG CAAACAAACA GATTAATCCG
 GAGCGCTGGG CTGATATTTT AGCCAAAGTC CGTGGGAAAA ATCTCCCTAA
 GGGCACAGAA GTACCCCGCT TGATTAAC CACAAGTCCT CCATCCTTGA
 GCCTAGAAGT CACACCACCT TTTCTGCTG TTTCTCCCC CTCAGCATCT
 CCTGTGCAGA CAGTAACCAG TGCTGAAGAA TCCTCAGCAG ATGTACCTCT
 ACTTGGTGAA GAAGAGCACG TTTTGGGTAC CATTTCTCA GCCAGCATGG
 GGCTAGAACA CAACCACAAT GGAGTTATC TTGTTGAACC TGAAGTAACA
 AGCACACCTC TGGAGGAAGT TGTTGATGAC CTTTCTGAGA AGACTGAGGA
 GATAACTTCC ACTGAAGGAG ACCTGAAGGG GACAGCAGCC CCTACACTTA
 TATCTGAGCC TTATGAACCA TCTCCTACTC TGCACACATT AGACACAGTC
 TATGAAAAGC CCACCCATGA AGAGACGGCA ACAGAGGGTT GGTCTGCAGC
 AGATGTTGGA TCGTCACCAG AGCCACATC CAGTGAGTAT GAGCCTCCAT
 TGGATGCTGT CTCCTTGGCT GAGTCTGAGC CCATGCAATA CTTTGACCCA
 GATTTGGAGA CTAAGTCACA ACCAGATGAG GATAAGATGA AAGAAGACAC
 CTTTGCACAC CTTACTCCAA CCCCCACCAT CTGGGTAAAT GACTCCAGTA
 CATCACAGTT ATTTGAGGAT TCTACTATAG GGGAACCAGG TGTCACAGGC
 CAATCACATC TACAAGGACT GACAGACAAC ATCCACCTTG TGAAAAGTAG
 TCTAAGCACT CAAGACACCT TACTGATTAA AAAGGGTATG AAAGAGATGT
 CTCAGACACT ACAGGGAGGA AATATGCTAG AGGGAGACCC CACACACTCC
 AGAAGTTCTG AGAGTGAGGG CCAAGAGAGC AAATCCATCA CTTTGCCTGA
 CTCCACACTG GGTATAATGA GCAGTATGTC TCCAGTTAAG AAGCCTGCGG
 AAACCACAGT TGGTACCCTC CTAGACAAAG ACACCACAAC AGTAACAACA
 ACACCAAGGC AAAAAGTTGC TCCGTCATCC ACCATGAGCA CTCACCCTTC
 TCGAAGGAGA CCAACGGGA GAAGGAGATT ACGCCCCAAC AAATTCCGCC
 ACCGGCACAA GCAAACCCCA CCCACAACCTT TTGCCCCATC AGAGACTTTT
 TCTACTCAAC CAACTCAAGC ACCTGACATT AAGATTTCAA GTCAAGTGGA
 GAGTTCTCTG GTTCCTACAG CTGGGTGGA TAACACAGTT AATACCCCA
 AACAGTTGGA AATGGAGAAG AATGCAGAAC CCACATCCAA GGGAACACCA
 CGGAGAAAAC ACGGGAAGAG GCCAAACAAA CATCGATATA CCCCTTCTAC
 AGTGAGCTCA AGAGCGTCCG GATCCAAGCC CAGCCCTTCT CCAGAAAATA
 AACATAGAAA CATTGTTACT CCCAGTTCAG AAATATACT TTTGCCTAGA
 ACTGTTTCTC TGAAAACCTGA GGGCCCTTAT GATTCCTTAG ATTACATGAC
 AACCACCAGA AAAATATATT CATCTTACCC TAAAGTCCAA GAGACACTTC
 CAGTCACATA

TAAACCCACA TCAGATGGAA AAGAAATTAA GGATGATGTT GCCACAAATG
 TTGACAAACA TAAAAGTGAC ATTTTAGTCA CTGGTGAATC AATTACTAAT
 GCCATACCAA CTTCTCGCTC CTTGGTCTCC ACTATGGGAG AATTTAAGGA
 AGAATCCTCT CCTGTAGGCT TTCCAGGAAC TCCAACCTGG AATCCCTCAA
 GGACGGCCCA GCCTGGGAGG CTACAGACAG ACATACCTGT TACCACTTCT
 GGGGAAAATC TTACAGACCC TCCCCTTCTT AAAGAGCTTG AGGATGTGGA
 TTCACTTCC GAGTTTTTGT CCTCTTTGAC AGTCTCCACA CCATTTCACC
 AGGAAGAAGC TGGTCTTCC ACAACTCTCT CAAGCATAAA AGTGGAGGTG
 GCTTCAAGTC AGGCAGAAAC CACCACCCTT GATCAAGATC ATCTTGAAAC
 CACTGTGGCT ATTCTCCTTT CTGAAACTAG ACCACAGAAT CACACCCCTA
 CTGCTGCCCCG GATGAAGGAG CCAGCATCCT CGTCCCCATC CACAATTCTC
 ATGTCTTTGG GACAAACCAC CACCACTAAG CCAGCACTTC CCAGTCCAAG
 AATATCTCAA GCATCTAGAG ATTCCAAGGA AAATGTTTTT TTGAATTATG
 TGGGGAATCC AGAAACAGAA GCAACCCCAG TCAACAATGA AGGAACACAG
 CATATGTCAG GGCCAAATGA ATTATCAACA CCCTCTTCCG ACCGGGATGC
 ATTTAACTTG TCTACAAAGC TGGAATTGGA AAAGCAAGTA TTTGGTAGTA
 GGAGTCTACC ACGTGGCCCA GATAGCCAAC GCCAGGATGG AAGAGTTCAT
 GCTTCTCATC AACTAACCAG AGTCCCTGCC AAACCCATCC TACCAACAGC
 AACAGTGAGG CTACCTGAAA TGTCCACACA AAGCGCTTCC AGATACTTTG
 TAACTTCCCA GTCACCTCGT CACTGGACCA ACAAACCGGA AATAACTACA
 TATCCTTCTG GGGCTTTGCC AGAGAACAAA CAGTTTACAA CTCCAAGATT
 ATCAAGTACA ACAATTCCTC TCCCATTGCA CATGTCCAAA CCCAGCATTC
 CTAGTAAGTT TACTGACCGA AGAACTGACC AATTCAATGG TTA CTCCAAA
 GTGTTTGGAA ATAACAACAT CCCTGAGGCA AGAAACCCAG TTGGAAAGCC
 TCCCAGTCCA AGAATTCCTC ATTATTCCAA TGGAAGACTC CCTTTCTTTA
 CCAACAAGAC TCTTTCTTTT CCACAGTTGG GAGTCACCCG GAGACCCAG
 ATACCCACTT CTCCTGCCCC AGTAATGAGA GAGAGAAAAG TTATTCCAGG
 TTCCTACAAC AGGATACATT CCCATAGCAC CTTCATCTG GACTTTGGCC
 CTCCGGCACC TCCGTTGTTG CACACTCCGC AGACCACGGG ATCACCTCA
 ACTAACTTAC AGAATATCCC TATGGTCTCT TCCACCCAGA GTTCTATCTC
 CTTTATAACA TCTTCTGTCC AGTCCTCAGG AAGCTTCCAC CAGAGCAGCT
 CAAAGTTCTT TGCAGGAGGA CCTCCTGCAT CCAAATTCTG GTCTCTTGGG
 GAAAAGCCCC AAATCCTCAC CAAGTCCCCA CAGACTGTGT CCGTCACCGC
 TGAGACAGAC ACTGTGTTCC CCTGTGAGGC AACAGGAAAA CCAAAGCCTT
 TCGTTACTTG GACAAAGGTT TCCACAGGAG CTCTTATGAC TCCGAATACC
 AGGATACAAC GGT TTGAGGT TCTCAAGAAC GGTACCTTAG TGATACGGAA
 GGTTCAGTA CAAGATCGAG GCCAGTATAT GTGCACCGCC AGCAACCTGC
 ACGGCCTGGA CAGGATGGTG GTCTTGCTTT CGGTCACCGT GCAGCAACCT
 CAAATCCTAG CCTCCCACTA CCAGGACGTC ACTGTCTACC TGGGAGACAC
 CATTGCAATG GAGTGTCTGG CCAAAGGGAC CCCAGCCCCC CAAATTTCCT
 GGATCTTCCC TGACAGGAGG GTGTGGCAAA CTGTGTCCCC

CGTGGAGAGC CGCATCACCC TGCACGAAAA CCGGACCCTT TCCATCAAGG
 AGGCGTCCTT CTCAGACAGA GCGTCTATA AGTGCGTGGC CAGCAATGCA
 GCCGGGGCGG ACAGCCTGGC CATCCGCCTG CACGTGGCGG CACTGCCCCC
 CGTTATCCAC CAGGAGAAGC TGGAGAACAT CTCGCTGCCC CCGGGGCTCA
 GCATTCACAT TCACTGCACT GCCAAGGCTG CGCCCCTGCC CAGCGTGCGC
 TGGGTGCTCG GGGACGGTAC CCAGATCCGC CCCTCGCAGT TCCTCCACGG
 GAACTTGTTT GTTTTCCCCA ACGGGACGCT CTACATCCGC AACCTCGCGC
 CCAAGGACAG CGGGCGCTAT GAGTGCCTGG CCGCCAACCT GGTAGGCTCC
 GCGCGCAGGA CGGTGCAGCT GAACGTGCAG CGTGCAGCAG CCAACGCGCG
 CATCACGGGC ACCTCCCCGC GGAGGACGGA CGTCAGGTAC GGAGGAACCC
 TCAAGCTGGA CTGCAGCGCC TCGGGGGACC CCTGGCCGCG CATCCTCTGG
 AGGCTGCCGT CCAAGAGGAT GATCGACGCG CTCTTCAGTT TTGATAGCAG
 AATCAAGGTG TTTGCCAATG GGACCCTGGT GGTGAAATCA GTGACGGACA
 AAGATGCCGG AGATTACCTG TGCCTAGCTC GAAATAAGGT TGGTGATGAC
 TACGTGGTGC TCAAAGTGGA TGTGGTGATG AAACCGGCCA AGATTGAACA
 CAAGGAGGAG AACGACCACA AAGTCTTCTA CGGGGGTGAC CTGAAAGTGG
 ACTGTGTGGC CACCGGGCTT CCCAATCCCG AGATCTCCTG GAGCCTCCCA
 GACGGGAGTC TGGTGAATC CTTCATGCAG TCGGATGACA GCGGTGGACG
 CACCAAGCGC TATGTCGTCT TCAACAATGG GACACTCTAC TTAAACGAAG
 TGGGGATGAG GGAGGAAGGA GACTACACCT GCTTTGCTGA AAATCAGGTC
 GGAAGGACG AGATGAGAGT CAGAGTCAAG GTGGTGACAG CGCCCGCCAC
 CATCCGGAAC AAGACTTACT TGGCGGTTCA GGTGCCCTAT GGAGACGTGG
 TCACTGTAGC CTGTGAGGCC AAAGGAGAAC CCATGCCCAA GGTGACTTGG
 TTGTCCCCAA CCAACAAGGT GATCCCCACC TCCTCTGAGA AGTATCAGAT
 ATACCAAGAT GGCCTCTCC TTATTAGAA AGCCAGCGT TCTGACAGCG
 GCAACTACAC CTGCCTGGTC AGGAACAGCG CGGGAGAGGA TAGGAAGACG
 GTGTGGATTG ACGTCAACGT CCAGCCACCC AAGATCAACG GTAACCCCAA
 CCCCATCACC ACCGTGCGGG AGATAGCAGC CGGGGGCAGT CGGAAACTGA
 TTGACTGCAA AGCTGAAGGC ATCCCCACCC CGAGGGTGTT ATGGGCTTTT
 CCCGAGGGTG TGGTTCTGCC AGCTCCATAC TATGGAAACC GGATCACTGT
 CCATGGCAAC GGTTCCTGG ACATCAGGAG TTTGAGGAAG AGCGACTCCG
 TCCAGCTGGT ATGCATGGCA CGCAACGAGG GAGGGGAGGC GAGGTTGATC
 GTGCAGCTCA CTGTCCTGGA GCCCATGGAG AAACCCATCT TCCACGACCC
 GATCAGCGAG AAGATCACGG CCATGGCGGG CCACACCATC AGCCTCAACT
 GCTCTGCCGC GGGGACCCCG ACACCCAGCC TGGTGTGGGT CCTTCCCAAT
 GGCACCGATC TGCAGAGTGG ACAGCAGCTG CAGCGCTTCT ACCACAAGGC
 TGACGGCATG CTACACATTA GCGGTCTCTC CTCGGTGGAC GCTGGGGCCT
 ACCGCTGCGT GGCCCGCAAT GCCGCTGGCC ACACGGAGAG GCTGGTCTCC
 CTGAAGGTGG GACTGAAGCC AGAAGCAAAC AAGCAGTATC ATAACCTGGT
 CAGCATCATC AATGGTGAGA CCCTGAAGCT CCCCTGCACC CCTCCCGGGG
 CTGGGCAGGG ACGTTTCTCC TGGACGCTCC CCAATGGCAT GCATCTGGAG

GGCCCCAAA CCCTGGGACG CGTTTCTCTT CTGGACAATG GCACCCTCAC
GGTTCGTGAG GCCTCGGTGT TTGACAGGGG TACCTATGTA TGCAGGATGG
AGACGGAGTA CGGCCCTTCG GTCACCAGCA TCCCCGTGAT TGTGATCGCC
TATCCTCCCC GGATCACCAG CGAGCCCACC CCGGTCATCT ACACCCGGCC
CGGGAACACC GTGAAACTGA ACTGCATGGC TATGGGGATT CCCAAAGCTG
ACATCACGTG GGAGTTACCG GATAAGTCGC ATCTGAAGGC AGGGGTTCAG
GCTCGTCTGT ATGGAAACAG ATTTCTTCAC CCCAGGGGAT CACTGACCAT
CCAGCATGCC ACACAGAGAG ATGCCGGCTT CTACAAGTGC ATGGCAAAAA
ACATTCTCGG CAGTGACTCC AAAACAACCTT ACATCCACGT CTTCTGAAAT
GTGGATTCCA GAATGATTGC TTAGGAACTG ACAACAAAGC GGGGTTTGT
AGGGAAGCCA GGTGGGGAA TAGGAGCTCT TAAATAATGT GTCACAGTGC
ATGGTGGCCT CTGGTGGGTT TCAAGTTGAG GTTGATCTTG ATCTACAATT
GTTGGGAAAA GGAAGCAATG CAGACACGAG AAGGAGGGCT CAGCCTTGCT
GAGACACTTT CTTTGTGTT TACATCATGC CAGGGGCTTC ATTCAGGGTG
TCTGTGCTCT GACTGCAATT TTTCTTCTTT TGCAAATGCC ACTCGACTGC
CTTCATAAGC GTCCATAGGA TATCTGAGGA ACATTATCA AAAATAAGCC
ATAGACATGA ACAACACCTC ACTACCCCAT TGAAGACGCA TCACCTAGTT
AACCTGCTGC AGTTTTTACA TGATAGACTT TGTTCCAGAT TGACAAGTCA
TCTTTCAGTT ATTTCTCTG TCACTTCAA ACTCCAGCTT GCCCAATAAG
GATTTAGAAC CAGAGTGACT GATATATATA TATATATTTT AATTCAGAGT
TACATACATA CAGCTACCAT TTTATATGAA AAAAGAAAAA CATTCTTCC
TGGAACCTCAC TTTTATATA ATGTTTTATA TATATATTTT TTCCTTTCAA
ATCAGACGAT GAGACTAGAA GGAGAAATAC TTTCTGTCTT ATTAAAATTA
ATAAATTATT GGTCTTTACA AGACTTGGAT ACATTACAGC AGACATGGAA
ATATAATTTT AAAAAATTC TCTCCAACCT CCTTCAAATT CAGTCACCAC
TGTTATATTA CCTTCTCCAG GAACCCTCCA GTGGGGAAGG CTGCGATATT
AGATTTCTT GTATGCAAAG TTTTGTGTA AAGCTGTGCT CAGAGGAGGT
GAGAGGAGAG GAAGGAGAAA ACTGCATCAT AACTTTACAG AATTGAATCT
AGAGTCTTCC CCGAAAAGCC CAGAACTTC TCTGCAGTAT CTGGCTTGTC
CATCTGGTCT AAGGTGGCTG CTTCTTCCCC AGCCATGAGT CAGTTTGTGC
CCATGAATAA TACACGACCT GTTATTTCCA TGACTGCTTT ACTGTATTTT
TAAGGTCAAT ATACTGTACA TTTGATAATA AAATAATATT CTCCCAAAAA AAAAA

Human OCP: nucleotide sequence of ORF

ATGAAGGTAAAAGGCAGAGGAATCACCTGCTTGCTGGTCTCCTTTGCTGTGATCTGC
CTGGTCGCCACCCCTGGGGGCAAGGCCTGTCTCGCCGCTGTGCCTGTTATATGCCT
ACGGAGGTACACTGCACATTTTCGGTACCTGACTTCCATCCCAGACAGCATCCCCGCC
AATGTGGAACGCATCAATTTAGGATACAACAGCTTGGTTAGATTGATGGAAACAGAT
TTTTCTGGCCTGACCAAACCTGGAGTTACTCATGCTTCACAGCAATGGCATTACACA
ATCCCTGACAAGACCTTCTCAGATTTGCAGGCCTTGCAGGTCTTAAAAATGAGCTAT
AATAAAGTCCGAAACCTTCAGAAAGATACTTTTTATGGCCTCAGGAGCTTGACACGA
TTGCACATGGACCACAACAATATTGAGTTTATAAACCCAGAGGTTTTTTATGGGCTC
AACTTTCTCCGCCTGGTGCCTTGGAAAGGAAATCAGCTCACTAAGCTCCACCCAGAT
ACATTTGTCTCTTTGAGCTACCTCCAGATATTTAAATCTCTTTCATTAAGTTCCTA
TACTTGTCTGATAACTTCCTGACCTCCCTCCCTCAAGAGATGGTCTCCTATATGCCT
GACCTAGACAGCCTTTACCTGCATGGAAACCCATGGACCTGTGATTGCCATTTAAAG
TGGTTGTCTGACTGGATACAGGAGAAGCCAGATGTAATAAAATGCAAAAAGATAGA
AGTCCCTCTAGTGCTCAGCAGTGTCCACTTTGCATGAACCCTAGGACTTCTAAAGGC
AAGCCGTTAGCTATGGTCTCAGCTGCAGCTTTCAGTGTGCCAAGCCAACCATTGAC
TCATCCCTGAAATCAAAGAGCCTGACTATTCTGGAAGACAGTAGTTCTGCTTTCATC
TCTCCCCAAGGTTTCATGGCACCCTTTGGCTCCCTCACTTTGAATATGACAGATCAG
TCTGGAAATGAAGCTAACATGGTCTGCAGTATTCAAAAGCCCTCAAGGACATCACCC
ATTGCATTCACTGAAGAAAATGACTACATCGTGCTAAATACTTCATTTTCAACATTT
TTGGTGTGCAACATAGATTACGGTCACATTCAGCCAGTGTGGCAAATTTTGGCTTTG
TACAGTGATTCTCCTCTGATACTAGAAAGGAGCCACTTGCTTAGTGAAACACCGCAG
CTCTATTACAAATATAAACAGGTGGCTCCTAAGCCTGAAGACATTTTACCAACATA
GAGGCAGATCTCAGAGCAGATCCCTCTTGGTTAATGCAAGACCAAATTTCTTGCAG
CTGAACAGAACTGCCACCACATTCAGTACATTACAGATCCAGTACTCCAGTGATGCT
CAAATCACTTTACCAAGAGCAGAGATGAGGCCAGTGAAACACAAATGGACTATGATT
TCAAGGGATAACAATACTAAGCTGGAACATACTGTCTTGGTAGGTGGAACCGTTGGC
CTGAACTGCCAGGCCAAGGAGACCCCAACCCACACGTGGATTGGCTTCTAGCTGAT
GGAAGTAAAGTGAGAGCCCCTTATGTGCTGAGGATGGACGGATCCTAATAGACAAA
AGTGGAATAATTGGAACCTCAGATGGCTGATAGTTTTTGACACAGGCGTATATCACTGT
ATAAGCAGCAATTATGATGATGCAGATATTCTCACCTATAGGATAACTGTGGTAGAA
CCTTTGGTCGAAGCCTATCAGGAAAATGGGATTTCATCACACAGTTTTTCATTGGTGAA
ACACTTGATCTTCCATGCCATTCTACTGGTATCCCAGATGCCTCTATTAGCTGGGTT
ATTCCAGGAAACAATGTGCTCTATCAGTCATCAAGAGACAAGAAAGTTCTAAACAAT
GGCACATTAAGAATATTACAGGTCACCCCGAAAGACCAAGGTTATTATCGCTGTGTG
GCAGCCAACCCATCAGGGGTTGATTTTTTTGATTTTCCAAGTTTCAGTCAAGATGAAA
GGACAAAGGCCCTTGGAGCATGATGGAGAAACAGAGGGATCTGGACTTGATGAGTCC
AATCCTATTGCTCATCTTAAGGAGCCACCAGGTGCACAACTCCGTACATCTGCTCTG
ATGGAGGCTGAGGTTGGAAAACACACCTCAAGCACAAAGTAAGAGGCACAACTATCGG
GAATTAACACTCCAGCGACGTGGAGATTCAACACATCGACGTTTTAGGGAGAATAGG
AGGCATTTCCCTCCCTCTGCTAGGAGAATTGACCCACAACATTGGGCGGCCTGTTG
GAGAAAGCTAAAAAGAATGCTATGCCAGACAAGCGAGAAAATACCACAGTGAGCCCA
CCCCAGTGGTCACCCAACTCCCAAACATACCTGGTGAAGAAGACGATTCTCAGGC
ATGCTCGCTCTACATGAGGAATTTATGGTCCCGGCCACTAAAGCTTTGAACCTTCCA
GCAAGGACAGTGACTGCTGACTCCAGAACAATATCTGATAGTCCTATGACAAACATA
AATTATGGCACAGAATTCTCTCCTGTTGTGAATTCACAAATACTACCACCTGAAGAA
CCCACAGATTTCAAACGTCTACTGCTATTAAACTACAGCCATGTCAAAGAATATA

AACCCCAACCATGTCAAGCCAAATACAAGGCACAACCAATCAACATTTCATCCACTGTC
TTTCCACTGCTACTTTGGAGCAACTGAATTTTCAGGACTCTGACCAGATGGGAAGAGGA
AGAGAGCATTTCCAAAGTAGACCCCCAATAACAGTAAGGACTATGATCAAAGATGTC
AATGTCAAAATGCTTAGTAGCACCACCAACAACTATTATTAGAGTCAGTAAATACC
ACAAATAGTCATCAGACATCTGTAAGAGAAGTGAGTGAACCCAGGCACAATCACTTC
TATTCTCACACTACTCAAATACTTAGCACCTCCACGTTCCCTTCAGATCCACACACA
GCTGCTCATTCTCAGTTTCCGATCCCTAGAAATAGTACAGTTAACATCCCGCTGTTC
AGACGCTTTGGGAGGCAGAGGAAAATTGGCGGAAGGGGGCGGATTATCAGCCCATAT
AGAACTCCAGTTCTGCGACGGCATAGATACAGCATTTTTCAGGTCAACAACCAGAGGT
TCTTCTGAAAAAAGCACTACTGCATTCTCAGCCACAGTGCTCAATGTGACATGTCTG
TCCTGTCTTCCCAGGGAGAGGCTCACCCTGCCACAGCAGCATTGTCTTTTCCAAGT
GCTGCTCCCATCACCTTCCCCAAAGCTGACATTGCTAGAGTCCCATCAGAAGAGTCT
ACAACTCTAGTCCAGAATCCACTATTACTACTTGAGAACAAACCCAGTGTAGAGAAA
ACAACACCCACAATAAAATATTTTCAGGACTGAAATTTCCCAAGTGACTCCAAGTGGT
GCAGTCATGACATATGCTCCAACATCCATACCCATGGAAAAAACTCACAAAGTAAAC
GCCAGTTACCCACGTGTGTCTAGCACCAATGAAGCTAAAAGAGATTTCAGTGATTACA
TCGTCACTTTTCAGGTGCTATCACCAAGCCACCAATGACTATTATAGCCATTACAAGG
TTTTCAAGAAGGAAAATTCCCTGGCAACAGAACTTTGTAAATAACCATAACCCAAAA
GGCAGATTAAAGGAATCAACATAAAGTTAGTTTACAAAAAAGCACAGCTGTGATGCTT
CCTAAAACATCTCCTGCTTTACCACAGAGACAAAGTTCCCTTTTCATTTTACCACA
CTTTCAACAAGTGTGATGCAAATTCCATCTAATACCTTGACTACCGCTCACCACACT
ACGACCAAAAACACACAATCCTGGAAGTCTTCCAACAAAGAAGGAGCTTCCCTTCCCA
CCCCTTAACCTTATGCTTCTTAGTATTATAAGCAAAGACTCAAGTACAAAAGCATC
ATATCAACGCAAACAGCAATACCAGCAACAACCTTACCTTCCCTGCATCTGTCTATC
ACTTATGAAACCCAAACAGAGAGATCTAGAGCACAAACAATACAAAGAGAACAGGAG
CCTCAAAAGAAGAACAGGACTGACCCAAACATCTCTCCAGACCAGAGTTCTGGCTTC
ACTACACCCACTGCTATGACACCTCCTGCTCTGGCATTCACTCATTCCCCACCAGAA
AACACAACCTGGGATTTCAAGCACAAATCAGTTTTTCATTCAAGAACTCTTAATCTGACA
GATGTGATTGAAGAACTAGCCCAAGCAAGTACTCAGACTTTGAAGAGCACAAATTGCT
TCTGAAACAACCTTTGTCCAGCAAATCACACCAGAGTACCACAACCTAGGAAAGCATCA
TTAGACACTCCCATACCACCATTCTTGAGCAGCAGTGCTACTCTAATGCCAGTTCCC
ATCTCCCCCTCCCTTTACTCAGAGAGCAGTTACTGACACACGTGGCGACTCCCATTTC
CGGCTTATGACAAATACAGTGGTCAAGCTGCACGAATCCTCAAGGCACAATCTCCAA
ATGCCAAGTTCACAATTGGAACCACTCACTTCATCTACCTCTAATCTGTTACATTCT
ACTCCCATGCCAGCACTAACAACAGTTAAATCACAGAATTCCAAATTAACCTCCATCT
CCCTGGGCAGAATACCAATTTTGGCACAACCATACTCAGACATTGCTGAAAAAGGC
AAAAAGCCAGAAGTAAGCATGTTGGCTACTACAGGCCTGTCCGAGGCCACCACTCTT
GTTTCAGATTGGGATGGACAGAAGAACACAAAGAAGAGTGACTTTGATAAGAAACCA
GTTCAAGAAGCAACAACCTTCCAACTCCTTCCCTTTGACTCTTTGTCTAGGTATATA
TTTGAAGAGCCAGGATAGTTGGAGGAAAAGCTGCAAGTTTTACTATTCCAGCTAAC
TCAGATGCCTTTCTTCCCTGTGAAGCTGTTGGAAATCCCCTGCCACCATTTCATTGG
ACCAGAGTTTCAGGACTTGATTTATCTAGAGGAAACCAGAATAGCAGGGTCCAGGTT
CTCCCCAATGGTACCCTGTCCATCCAGAGGGTGGAAATTCAGGACCGCGGACAGTAC
TTGTGTTCCGCATCCAATCTGTTTGGCACAGACCACCTTCATGTACCTTGTCTGTG
GTTTCCTATCCTCCCAGGATCCTGGAGAGACGTACCAAAGAGATCACAGTTCATTCC
GGAAGCACTGTGGAAGTGAAGTGCAGAGCAGAAGGTAGGCCAAGCCCTACAGTTACC
TGGATTCTTGCAAACCAAACAGTTGTCTCAGAATCATCCCAGGGAAGTAGGCAGGCT
GTGGTGACGGTTGACGGAACATTGGTCCTCCACAATCTCAGTATTTATGACCGTGGC
TTTTACAAATGTGTGGCCAGCAACCCAGGTGGCCAGGATTCACTGCTGGTTAAAATA
CAAGTCATTGCAGACCACTGTTATTCTAGAGCAAAGGAGGCAAGTCATTGTAGGC
ACTTGGGGTGAAAGTTTAAAACCTGCCCTGTACTGCAAAAGGAACTCCTCAGCCCAGC

GTTTACTGGGTCCTCTCTGATGGCACTGAAGTGAAACCATTACAGTTTACCAATTCC
AAGTTGTTCTTATTTTCAAATGGGACTTTGTATATAAGAAACCTAGCCTCTTCAGAC
AGGGGCACTTATGAATGCATTGCTACCAGTTCCACTGGTTTCGGAGCGAAGAGTAGTA
ATGCTTACAATGGAAGAGCGAGTGACCAGCCCCAGGATAGAAGCTGCATCCCAGAAA
AGGACTGAAGTGAATTTTGGGGACAAATTACTACTGAACTGCTCAGCCACTGGGGAG
CCCAAACCCCCAAATAATGTGGAGGTTACCATCCAAGGCTGTGGTTCGACCAGTGGAGC
TGGATCCACGTCTACCCTAATGGATCCCTGTTTATTGGATCAGTAACAGAAAAAGAC
AGTGGTGTCTACTTGTGTGTGGCAAGAAACAAAATGGGGGATGATCTGATACTGATG
CATGTTAGCCTAAGACTGAAACCTGCCAAAATTGACCACAAGCAGTATTTTAGAAAG
CAAGTGCTCCATGGGAAAGATTTCCAAGTAGATTGCAAAGCTTCCGGCTCCCCAGTG
CCAGAGATATCTTGGAGTTTGCCTGATGGAACCATGATCAACAATGCAATGCAAGCC
GATGACAGTGGCCACAGGACTAGGAGATATACCCTTTTCAACAATGGAACCTTTATAC
TTCAACAAAGTTGGGGTAGCGGAGGAAGGAGATTATACTTGCTATGCCCAGAACACC
CTAGGGAAAGATGAAATGAAGGTCCACTTAACAGTTATAACAGCTGCTCCCCGGATA
AGGCAGAGTAACAAAACCAACAAGAGAATCAAAGCTGGAGACACAGCTGTCCTTGAC
TGTGAGGTCACTGGGGATCCCAAACCAAAAATATTTTGGTTGCTGCCTTCCAATGAC
ATGATTTTCTTCTCCATTGATAGGTACACATTTTCATGCCAATGGGTCTTTGACCATC
AACAAAGTGAACTGCTCGATTCTGGAGAGTACGTATGTGTAGCCCGAAATCCCAGT
GGGGATGACACCAAAATGTACAACTGGATGTGGTCTCTAAACCTCCATTAATCAAT
GGTCTGTATACAAACAGAACTGTTATTAAAGCCACAGCTGTGAGACATTCCAAAAAA
CACTTTGACTGCAGAGCTGAAGGGACACCATCTCCTGAAGTCATGTGGATCATGCCA
GACAATATTTTCTCACAGCCCCATACTATGGAAGCAGAATCACAGTCCATAAAAAT
GGAACCTTGGAATTAGGAATGTGAGGCTTTCAGATTCAGCCGACTTTATCTGTGTG
GCCCCAAATGAAGGTGGAGAGAGCGTGTTGGTAGTACAGTTAGAAGTACTGGAATG
CTGAGAAGACCGACATTTAGAAATCCATTTAATGAAAAAATAGTTGCCCAGCTGGGA
AAGTCCACAGCATTGAATTGCTCTGTTGATGGTAACCCACCACCTGAAATAATCTGG
ATTTTACCAAATGGCACACGATTTTCCAATGGACCACAAAGTTATCAGTATCTGATA
GCAAGCAATGGTTCTTTTATCATTTCTAAAACAACCTCGGGAGGATGCAGGAAAATAT
CGCTGTGCAGCTAGGAATAAAGTTGGCTATATTGAGAAATTAGTCATATTAGAAATT
GGCCAGAAGCCAGTTATTCTTACCTATGCACCAGGGACAGTAAAAGGCATCAGTGGA
GAATCTCTATCACTGCATTGTGTGTCTGATGGAATCCCTAAGCCAAATATCAAATGG
ACTATGCCAAGTGGTTATGTAGTAGACAGGCCCTCAAATTAATGGGAAATACATATTG
CATGACAATGGCACCTTAGTCATTAAAGAAGCAACAGCTTATGACAGAGGAACTAT
ATCTGTAAGGCTCAAATAGTGTTGGTCATACACTGATTACTGTTCCAGTAATGATT
GTAGCCTACCCTCCCCGAATTACAAATCGTCCACCCAGGAGTATTGTCACCAGGACA
GGGGCAGCCTTTCAGCTCCACTGTGTGGCCTTGGGAGTTCCCAAGCCAGAAATCACA
TGGGAGATGCCTGACCACTCCCTTCTCTCAACGGCAAGTAAAGAGAGGACACATGGA
AGTGAGCAGCTTCACTTACAAGGTACCCTAGTCATTCAGAATCCCCAAACCTCCGAT
TCTGGGATATACAAATGCACAGCAAAGAACCCACTTGGTAGTGATTATGCAGCAACC
TATATTCAAGTAATCTGA

Human OCP: predicted amino acid sequence

MKVKGRGITC LLVSFAVICL VATPGGKACP RRCACYMPTE VHCTFRYLTS
 IPDSIPPNVE RINLGYNLV RLMETDFSGL TKLELLMLHS NGIHTIPDKT
 FSDLQALQVL KMSYNKVRKL QKDTFYGLRS LTRLHMDHNN IEFINPEVFY
 GLNFLRLVHL EGNQLTKLHP DTFVSLSYLQ IFKISFIKFL YLSDNFLTSL
 PQEMVSYMPD LDSLYLHGNP WTCDCHLKWL SDWIQEKPDV IKCKKDRSPS
 SAQQCPLCMN PRTSKGKPLA MVSAAAFQCA KPTIDSSLKS KSLTILEDSS
 SAFISPOGFM APFGSLTLNM TDQSGNEANM VCSIQKPSRT SPIAFTEEND
 YIVLNTSFST FLVCNIDYGH IQPVWQILAL YSDSPLILER SHLLSETPQL
 YYKYKQVAPK PEDIFTNIEA DLRADPSWLM QDQISLQLNR TATTFSTLQI
 QYSSDAQITL PRAEMRPVKH KWTMISRDNN TKLEHTVLVG GTVGLNCPGQ
 GDPTPHVDWL LADGSKVRAP YVSEDGRILI DKSGKLELQM ADSFDTGVYH
 CISSNYDDAD ILTYRITVVE PLVEAYQENG IHHTVFIGET LDLPCHSTGI
 PDASISWVIP GNNVLYQSSR DKKVLNNGTL RILQVTPKDQ GYYRCVAANP
 SGVDFLIQV SVKMKGQRPL EHDGETEGSG LDESNPIAHL KEPPGAQLRT
 SALMEAEVGK HTSSTSKRHN YRELTQLRRG DSTHRRFREN RRHFPPSARR
 IDPQHWAAALL EKAKKNAMPD KRENTTVSPP PVVTQLPNIP GEEDDSSGML
 ALHEEFMVPA TKALNLPART VTADSRTISD SPMTNINYGT EFSPVVNSQI
 LPPEEPTDFK LSTAIKTAM SKNINPTMSS QIQGTTNQHS STVFPLLLGA
 TEFQDSQMG RGREHFQSRP PITVRTMIKD VNVKMLSST NKLLLESVNT
 TNSHQTSVRE VSEPRHNFY SHTTQILSTS TFPSPHTAA HSQFPIPRNS
 TVNIPLFRRF GRQRKIGGRG RIISPYRTPV LRRHRYSI FR STTRGSSEKS
 TTAFSATVLN VTCLSCLPRE RLTTATAALS FPSAAPITFP KADIARVPSE
 ESTTLVQNPL LLENKPSVE KTTPTIKYFR TEISQVTPTG AVMTYAPTISI
 PMEKTHKVNA SYPRVSSNE AKRDSVITSS LSGAITKPPM TIIAITRFSR
 RKIPWQQNFV NNHNPKGRLR NQHKVSLQKS TAVMLPKTSP ALPQRQSSPF
 HFTTLSTSVM QIPSNLTITA HHTTTKTHNP GSLPTKKELP FPPLNPMPLPS
 IISKDSSTKS IISTQTAIPA TTPTFPASVI TYETQTERS AQTIQREQEP
 QKKNRTDPNI SPDQSSGFTT PTAMTPPALA FTHSPPENTT GISSTISFHS
 RTLNLTDVIE ELAQASTQTL KSTIASSETTL SSKSHQSTTT RKASLDTPIP
 PFLSSSATLM PVPISPPFTQ RAVTDTRGDS HFRLMTNTVV KLHESSRHNL
 QMPSSQLEPL TSSTSNNLLHS TMPALTTVK SQNSKLTSP WAEYQFWHKP
 YSDIAEKGGK PEVSMLATTG LSEATTLVSD WDGQKNTKKS DFDKKPVQEA
 TTSKLLPFDS LSRYIFEKPR IVGGKAASFT IPANSDAFLP CEAVGNPLPT
 IHWTRVSGLD LSRGNQNSRV QVLPNGTSLI QRVEIQDRGQ YLCSASNLF
 TDHLHVTLSV VSYPPRILER RTKEITVHSG STVELKCRAE GRPSPTVTWI
 LANQTVVSES SQGSRQAVVT VDGTLVLHNL SIYDRGFYKC VASNPGGQDS
 LLVKIQVIAA PPVILEQRRQ VIVGTWGESL KLPCTAKGTP QPSVYVWLSD
 GTEVKPLQFT NSKLFLFSNG TLYIRNLASS DRGTYECIAT SSTGSERRVV
 MLTMEERVTS PRIEASQKR TEVNFGDKLL LNCSATGEPK PQIMWRLPSK
 AVVDQWSWIH VYPNGSLFIG SVTEKDSGVY LCVARNKMGD DLILMHVSLR
 LKPAKIDHKQ YFRKQVLHGK DFQVDCASG SPVPEISWSL PDGTMINNAM
 QADDSGHRTR RYTLFNNGTL YFNKVGVAEE GDYTCYAQNT LGKDEMKVHL
 TVITAAPRIR QSNKTNKRIK AGDTAVLDCE VTGDPKPKIF WLLPSNDMIS
 FSIDRYTFHA NGSLTINKVK LLDSGEYVCV ARNPSGDDTK MYKLDVVSKEP
 PLINGLYTNR TVIKATAVRH SKKHFDCAE GTPSPEVMWI MPDNIFLTAP
 YYGSRTIVHK NGTLEIRNVR LSDSADFICV ARNEGGEVL VVQLEVLEML
 RRPTFRNPFN EKIVAQLGKS TALNCSVDGN PPPEIIWILP NGTRFSNGPQ
 SYQYLIASNG SFIISKTTRE DAGKYRCAAR NKVGYIEKLV ILEIGQKPMI

LTYAPGTVKG ISGESLSLHC VSDGIPKPNI KWTMPSGYVV DRPQINGKYI
LHDNGTLVIK EATAYDRGNY ICKAQNSVGH TLITVPVMIV AYPPRITNRP
PRSIVTRTGA AFQLHCVALG VPKPEITWEM PDHSLSTAS KERTHGSEQL
HLQGTLVION PQTSDSGIYK CTAKNPLGSD YAATYIQVI*

SECRET